EXPLORING THE EFFECTS OF WORKING MEMORY CAPACITY, ATTENTION, AND EXPERTISE ON SITUATION AWARENESS IN A FLIGHT SIMULATION ENVIRONMENT

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

ORÇUN ORKAN ÖZCAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE
OF MASTER OF SCIENCE
IN
DEPARTMENT OF COGNITIVE SCIENCE

SEPTEMBER 2012

EXPLORING THE EFFECTS OF EXPERTISE AND WORKING MEMORY CAPACITY ON SITUATION AWARENESS IN FLIGHT SIMULATOR GAME ENVIRONMENTS

Submitted by **ORÇUN ORKAN ÖZCAN** in partial fulfillment of the requirements for the degree of **Master of Science in Cognitive Science, Middle Eact Technical University** by,

Prof. Dr. Nazife Baykal ____________________
Director, **Informatics Institute**

Prof. Dr. Cem Bozşahin ____________________
Head of Department**, Cognitive Science**

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

Dr. Bilge Say ____________________
Co-Supervisor, **Organizational Development and Planning Office, METU**

Examining Committee Members:

Prof. Dr. Cem Bozşahin ____________________
Cognitive Science, METU

Assist. Prof. Dr. Murat Perit Çakır ____________________
Health Informatics, METU

Dr. Bilge Say ____________________
Organizational Development and Planning Office, METU

Assist. Prof. Dr. Cengiz Acartürk ____________________
Cognitive Science, METU

Prof. Dr. Kürşat Çağıltay ____________________
Computer Education and Instructional Technology, METU

Date: 13/09/2012

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I have fully cited and referenced all materials and results that are not original to this work.

Name, Last name : Orçun Orkan Özcan

Signature : _________________

ABSTRACT

EXPLORING THE EFFECTS OF WORKING MEMORY CAPACITY, ATTENTION, AND EXPERTISE ON SITUATION AWARENESS IN A FLIGHT SIMULATION ENVIRONMENT

Özcan, Orçun Orkan
Master of Science, Department of Cognitive Science
Supervisor: Assist. Prof. Dr. Murat Perit Çakır
Co-Supervisor: Dr. Bilge Say

September 2012, 114 Pages

Experienced and novice simulator pilots are subject to some of the constraints of a real flight (or a full flight simulator) situation in a PC based flight simulation. In this thesis, the effects of expertise, working memory capacity, inhibition and divided attention on situation awareness (SA) in simulated flight task environments are investigated. The cognitive aspects underlying the process of situation awareness are explored by analyzing the compound effects of above listed factors. Online and Offline SA measurements obtained from a simulated flight task are used with flight hours standing for expertise and scores of Automated Operation Span Task, Stroop and Coşkunöz visual attention tasks as measurements for working memory capacity, inhibition and divided attention respectively. Regression analyses reveal that expected relationships of simulator pilots' SA with expertise and inhibition capacity are supported. On the other hand, expected relationships of SA with working memory and divided attention capacities are not revealed. This situation probably results from unsystematic differences in simulator pilots' practices.

In addition to the main experiment, simulator pilots' levels of neural activity at their dorsolateral prefrontal cortex are also measured during their behavioral performance. The relationships among neural correlates of mental workload induced by the simulated flight are investigated by the functional near-infrared (fNIR) spectroscopy optical brain imaging technology. Theorized mental workload distinction in the scenario of the simulated flight task is successfully observed in both perceived workload declarations and oxygenation measurements by fNIR.

Keywords: situation awareness, working memory, attention, mental workload, fNIR

ÖZ

BİR UÇUŞ SİMULASYONU ORTAMINDA ÇALIŞMA BELLEĞİ KAPASİTESİ, DİKKAT VE DENEYİMİN DURUMSAL FARKINDALIK ÜZERİNDEKİ ETKİLERİNİN ARAŞTIRILMASI

Özcan, Orçun Orkan
Master, Bilişsel Bilimler Bölümü
Tez Yöneticisi: Yrd. Doç Dr. Murat Perit Çakır
Ortak Tez Yöneticisi: Dr. Bilge Say

Eylül 2012, 114 Sayfa

Deneyimli ve başlangıç seviyesi simülatör pilotları, PC temelli uçuş simülasyonunda gerçek uçuş (veya tam uçuş simulatörü) şartlarına benzer durumlara maruz kalırlar. Bu tezde, simüle uçuş görev ortamında deneyim, çalışma belleği, baskılama ve bölünmüş dikkat kapasitesinin durumsal farkındalık (DF) üzerindeki etkileri incelenmiştir. Durumsal farkındalık kavramının bilişsel temelleri, yukarıda belirtilen yeteneklerin birleşik etkileri analiz edilerek araştırılmıştır. Uçuş simülasyonu görevi ile elde edilen çevrimiçi ve çevrimdışı DF ölçümleri; deneyim için toplam uçuş saati bilgisi, çalışma belleği için Otomatize İşlem Erim Görevi, baskılama için Stroop ve bölünmüş dikkat kapasitesi için Coşkunöz Görsel Dikkat testi ölçümleri kullanılmıştır. Regresyon analizleri sonucunda deneyim ve baskılama kapasitesinin durumsal farkındalık ile beklenen ilişkileri görülmüştür. Diğer yandan, durumsal farkıdalığın çalışma belleği ve bölünmüş dikkat kapasitesiyle beklenen ilişkileri gözlemlenememiştir. Bu durum muhtemelen simülatör pilotlarının kişisel uygulamalarındaki sistematik olmayan farklılıklardan kaynaklanmaktadır.

Ana deneylere ek olarak, simülatör pilotlarının dorsolateral prefrontal korteksteki sinirsel aktivite davranışsal performanslar süresince ölçülmüştür. Simule uçuş, görevinin oluşturduğu bilişsel işyükünün sinirsel bağlantıları fonksiyonel Yakın Kızılötesi (fNIR) Spektroskopi optik beyin görüntüleme tekniği ile incelenmiştir. Simule uçuş görevi senaryosunda kurgulanan bilişsel işyükü ayrımı hem algılanan işyükü bildirimlerinde hem de fNIR ile ölçülen oksijenlenmede gözlemlenmiştir.

Anahtar kelimeler: durumsal farkındalık, çalışma belleği, dikkat, bilişsel işyükü, fNIR

DEDICATION

To my family

ACKNOWLEDGMENTS

I would like to heartily thank to my supervisors, Assist. Prof. Dr. Murat Perit Çakır and Dr. Bilge Say. Without their support, this study would not be successfully completed. I also would like to express my gratitude to my previous supervisor Assist. Prof. Dr. Mine Mısırlısoy who had important contributions in the early periods of this thesis.

I want to acknowledge and thank to METU MODSIMMER for enabling me to use fNIRS systems, Subject Matter Experts for their help in the modification of the flight scenario and THK Antalya Branch members for their support in finding participants.

I am also grateful to Prof. Dr. Canan Sumer and her team in METU Psychology Department and Coşkunöz Metal Forming Company for their consent of use of divided attention test.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Situation Awareness (SA), also known as Situational Awareness, is an operational term originating from the aviation domain, which is used for the pilot's (or operator's in general) knowledge about the environment and procedures. This knowledge includes current information from the operation environment and the systems assisting the operator, task requirements and procedures, and foresight about the progress of the ongoing operation, briefly all information needed from the pilot. Endsley (1995b) reported that the 88% of all human errors are caused by inadequate SA. Relatedly, consequences of inadequate SA are important factors of accidents. Among these, perceptual factors caused 80.2% of the accidents, failure in comprehension caused 16.9% of them, and wrong predictions and decisions caused 2.3% of the accidents (Jones and Endsley, 1996). In this respect, operator's SA is regarded as a central concept in mission-critical tasks since it is strictly tied to operators' abilities and expected performance. Despite its critical role, there are inconsistent references to this term, lacking a commonly accepted definition.

The construct of SA is a quite rigorous term standing for composition of many psychological abilities. Due to its broad conceptualization, investigators of previous studies do not agree on a common definition. Mostly accepted conceptualizations of SA include not only grasping the environmental information but also envisioning about future determined by mission requirements. Considering the cognitive aspects of these conceptualizations, it can be seen that SA stands for a collection of diverse abilities like perception, long-term memory, working memory, attention, reasoning, learning, and decision-making (Sohn and Doane, 2004; Johannsdottir, 2004; McCarley, Wickens, Goh, and Horrey, 2002; Breton and Rousseau, 2001; Sukthankar, 1997; Endsley, 1997). The current status in the literature on SA does not cover a well-defined combination of abilities and there are inconsistencies in the way these terms are used. This situation motivates a need for cognitive elaboration in order to clarify SA's components as a complex cognitive phenomenon.

The most common definition of SA is given by Endsley (1995b) as a three level process: perceiving the environmental elements under time and space considerations, understanding these elements and estimating their future status. This definition is moderately accepted in various research approaches and reproduced by many authors around the same concept. To mention a few, Sarter and Woods (1991) define SA as the accessibility of an inclusive and clear picture of the situation refreshed dynamically with the outcomes of recurrent situation assessments. Fracker (1991) introduces a psychological term "working memory" into the definition functioning as storage and update mechanism for the existing knowledge. Main formulations of SA are based on studies with task-oriented approach. Tasks which SA is employed for, affect the overall conceptualization of this term. Consequently, definitions for SA are constructed as stand-alone functional descriptions which are rarely based on any fundamental cognitive concepts like working memory or attention. Endsley's (1995b) conceptualization is a relatively shallow analysis on SA in this respect, since it does not investigate the details of cognitive activities mentioned in her definition: perception, comprehension and projection. SA literature is mostly focused on conceptualization and measurement of the concept rather than analysis of underlying psychological properties.

Given the current situation of the dominant research paradigms, cognitive grounding of SA requires the uncovering of underlying processes. Cognitive components of SA were described by a limited number of studies (Durso & Gronlund, 1999; Endsley, 1995b; Sarter & Woods, 1991; Wickens, 1999). Moreover, individual differences in SA memory requirements is studied by a few researchers (Carretta, Perry & Ree, 1996; Johannsdottir, 2004; Sohn & Doane, 2004), details of which are given in the following paragraph. However, the experimental designs of these studies either lack an explicit assessment of SA in the task environment or do not involve operators/ pilots. Consequently, further explorations are needed for the cognitive grounding of SA. This study will try to explore the compound effects of working memory, attention mechanisms (inhibition and divided attention) and expertise in constructing SA.

In the literature, there are few studies using individual cognitive differences for the explanation of SA. Caretta, Perry and Ree (1996) stated that after controlling the effects of flight hours, general cognitive abilities composed of working memory, spatial reasoning and divided attention were found to be predictors of SA. However in this study, there is no information about these predictors' relative contribution to the explanation of SA and their effects were not clearly defined. Later, Johannsdottir (2004) had an important study for this area of research. It has been experimentally shown that working memory (WM) takes an active role in maintaining SA in a dynamic environment. One drawback of Johannsdottir's study is that experiments were carried out in the lab environment, not in a real task environment. Another remarkable study is by Sohn and Doane (2004) where they have showed the effects of long term and working memory on novice and expert groups. However, this study suffers from ecological validity since pilots took the SA tests not in the task environment (cockpit) but in office conditions. Also, simplistic SA tests were used instead of complete flights.

Çak's study (2011) improved this line of research and it is used as a baseline for the current study. In Çak's study, SA measurements were carried out with professional pilots at full flight simulators, investigating connections between SA and individual cognitive differences. Expertise, working memory, inhibition and divided attention capacities were used as predictors of SA performance. Significant relationships have been found between measures of SA and individual cognitive differences. For online SA Reaction Time (RT) measures, where participants answered the questions as the simulation runs, expertise, inhibition and divided attention capacity were found to be predictive. For offline SA measures, where participants answered questions on a paper after the simulation was frozen, expertise and working memory capacity explained the most of the variance in the SA scores. Current study aims to replicate the same experiment not on professional pilots but with simulator pilots. Simulator pilots, similar to professional pilots, have virtual (online) airline organizations where they perform online flights with other simulator pilots and virtual Air Traffic

Controllers. Consequently, Çak's experiments with minor modifications have been carried out with simulator pilots. Simulator pilots' SA have been measured and compared with their expertise level and inhibition, divided attention and working memory capacity. Similar to Çak's study, possible connections between SA, expertise and cognitive differences have been investigated to find significant relations between them.

Together with the behavioral experiments explained above, a brain imaging study with functional near-infrared (fNIR) spectroscopy has also been performed. Participants' brain activities from dorsolateral prefrontal cortex have been recorded to observe related activities during SA measurements. Different stages of the simulated flight task were compared to each other in order to investigate neural correlates of mental workload induced by different flight conditions.

At the end of behavioral experiments, meaningful relations between individual cognitive differences and SA scores are expected to be observed. Besides experience, working memory and divided attention capacity are candidates for positive predictors of SA scores and negative predictors for SA RT measurements. Inhibition delay, on the other hand, is a candidate of negative predictor of SA scores and positive predictor of SA RT measurements. Workload survey and brain activations recorded by fNIR are expected to reveal the relative changes in workload during the behavioral experiments.

In the following chapter, related literature on SA and individual cognitive differences is given. In Chapter 3, experiments carried out are described in detail. Results and discussions on the experiments, followed by the conclusions are presented in Chapter 4.

CHAPTER 2

LITERATURE REVIEW

In the following chapter, first literature on SA and individual differences such as working memory, attention and expertise is given. Second, research regarding the relationship between SA and individual differences is reviewed. Lastly, relevant studies on workload assessment using the optical brain imaging technique, fNIRS, are summarized within the scope of this study.

2.1 Situation Awareness

In this section, the relevant parts of the SA literature are reviewed and they will be given below under two section headings. First, different conceptualizations of SA will be given with exemplar models. SA measurements will be summarized in the last section.

2.1.1 Different Conceptualizations of Situation Awareness

Situation Awareness is an operational term that originated from the military domain for operations under critical time constraints and high mental workload. In a comprehensive review, Breton and Rousseau (2001) conclude that studies of SA have different viewpoints beginning from the definition of the concept. A classification of SA definitions based on these viewpoints is important for observing the tacit assumptions that the researchers hold.

Breton and Rousseau's review (2001) reveals that there is a fundamental difference in definitions of SA as to whether it is a state or a process. A prominent study by Endsley (1988) advocates the idea that SA is a state, which is gained by other processes embraced by the term situation assessment. In Endsley's approach, SA departs from all underlying activities and exhibits a mental incidence of the real situation of the environment at a given time. In contrast, researchers like Sarter and Woods (1995) and Adams, Tenney and Pew (1995) advocated for a different view. They asserted that SA is a "buzzword" and stands for a range of cognitive processes that jointly perform in the process. Similarly, Smith and Hancock (1995) claimed that SA could not be referred to as a form of declaration of conscious mental content at a given moment. They suggested that SA was the conceptualization of the ability to produce competent performance by appropriately directing consciousness in a dynamic environment.

As explained above, the main distinction in the definitions of SA is the state-process distinction. Two different approaches, situation-focused and operator-focused, led to this conceptual distinction as pointed by Durso and Gronlund (1999). Situation-focused approach employs state based explanations of SA and looks for a matching of information in the environment to the current mental state. For example, pilots' SA processes are defined in accordance with task sequences and requirements in a functionalist way. Consequently, this genre of definitions of SA inherits a task analysis structure. They are mainly proposals for successful operational designs of SA and they do not represent what really goes on psychologically. Operator-focused approach, on the other hand, deals with the identification of cognitive activities that the operator employs during the process of SA. This approach builds SA around the cognitive processes occurring in the operator's mind. However, cognitive processes are given very broadly like the perception/action cycle (Adams et al., 1995) without connections to cognitive psychological concepts like working memory or attention. Difference between the two approaches is also reflected in the models developed for SA. Below two models of SA are reviewed to summarize the main trends in the literature. Both of the models lack a cognitive perspective since they are not adequately integrated with experimental findings and their connections with cognitive states or processes are not investigated.

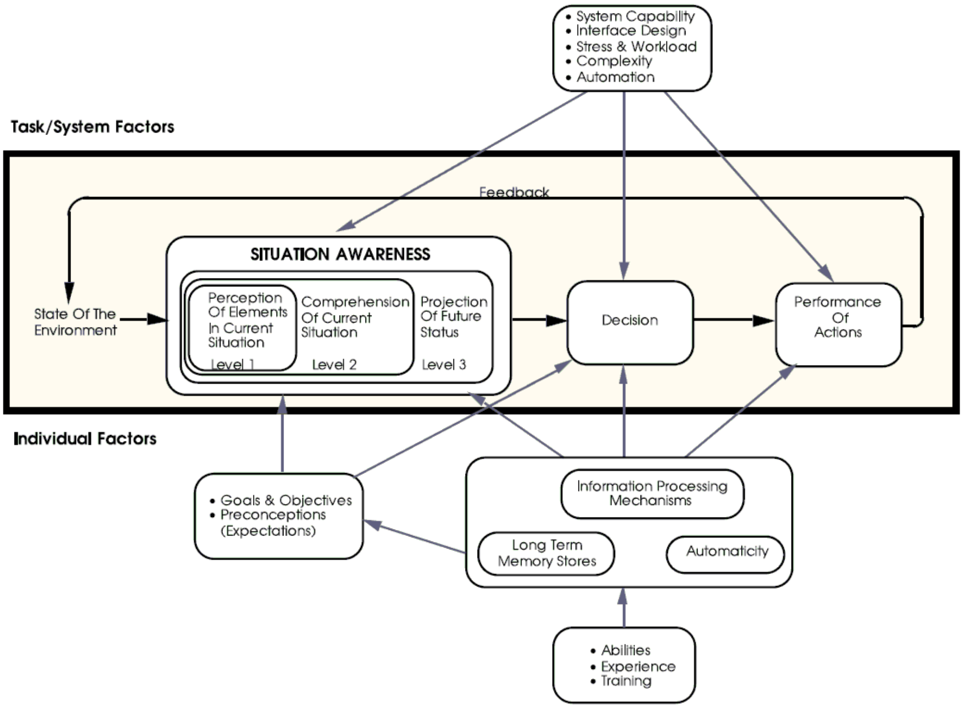


Figure 2.1: SA Model adopted from Endsley (1995b, p.35)

A representative model for the situation-focus approach is Endsley's model (1995b) given in Figure 2.1. It can be regarded as the most established model in the literature since it is elaborated and cited by many other studies on SA.

Endsley (2004) gives the "key factors" of the model as follows:

- Three levels of SA are formed by perception, comprehension and projection.
- Goals and goal-directed processing are important for attention control and information prioritization.
- Salient information can grab attention; however goal-driven processing can still be shifted back.
- Expectations take part in attention control and information prioritization.

- SA is limited by working memory demands especially for novices and for novel situations but these limits are pushed by mental models and pattern matching to prototypical schema.
- Mental models enable information synthesis and interpretation for future estimation.
- Situation based pattern-matching to schema for action selection.

In this model, three levels of operator activities are responsible for SA. Level 1 is perception of elements in the environment, which is the first step of information flow from the environment to the operator. In this step, the required information is extracted from the physical stimuli and internalized by the operator with the five senses. Second step, Level-2, is determined as comprehension of the current situation, which is the level where the operator evaluates the current situation by the combination and interpretation of all relevant information from Level 1 with goals and objectives. Level 3 is defined as projection of the future status and is responsible for foreseeing the future situation. The operator evaluates all the information and generates a scenario of the upcoming minutes in his/her mind to take action in the next level. Later stages of decision making are not an integral part of SA in this model.

The operator-focused approach, on the other hand, can be exemplified by the perception/action model (Adam et al., 1995). Their model, given in Figure2.2, is composed of three parts interacting with each other; the object in the environment, the schema in the operator’s mind and the exploration as an action. The notion of schema can be broadly described as a mental structure of ideas, an outline representing some aspect of the world. They also assert that schema (in the operator’s mind) is composed of explicit and implicit focus where explicit focus is matched to working memory and implicit focus deals with the inner working of the schema.

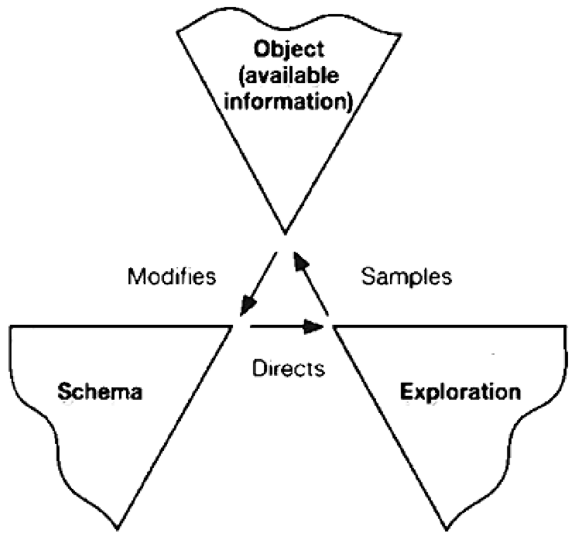


Figure 2.2: The perception/action cycle adopted from Adams et al. (1995, p.89)

To compare the two approaches, it can be said that situation-focused approach has a presumption of a clear picture of the situation. The situation in the environment is expected to be represented as a mental state. This conceptualization is also beneficial for SA models using information processing accounts, since it integrates all possible information into the notion of situation. On the contrary, operator-focused approach does not work on any well-defined aspect of situation, but focuses on the interactions between the operator and the environment. Unlike the information processing approach, the processes in the operator's mind are mentioned in connection with the interactions with the environment.

Given the different conceptualizations of SA, this study takes a closer standpoint to operator-based approaches. The effects of individual differences and expertise, which are underestimated in the situation-focused approach, will be centrally used in explaining SA based on its relations to those factors. However, none of the existing models of SA is directly employed in this thesis to examine cognitive underpinnings of SA, since overt experimental evidence and explanatory power on the models are lacking.

2.1.2 Situation Awareness Measurement

Similar to the conceptualization of SA and its models, SA measurement is also a controversial issue in the literature. Several researchers (Franker, 1991; Sarter & Woods, 1995) classified SA measurements in three general categories: (a) explicit, (b) implicit and (c) subjective measures of SA. Explicit measures are obtained by asking operators questions about the recent situation. Operators are expected to recall the SA contents relevant to the task (e.g. the airspeed of the plane for pilots, number of aircrafts waiting to land for air traffic controllers, etc.). Implicit measures of SA use task performance in order to guess SA. Performance values can be used as implicit measures to be evaluated for inferring SA. Third class in the SA measurement is subjective assessment. This method uses assessment ratings for operators or observers. Observers or operators themselves rate their personal opinion about the level of SA the operator has reached. For the scope of this study, only two explicit measures of SA are reviewed. These measures are selected to be used since they are commonly used in the literature and they measure different aspects of SA.

In an extensive review, Salmon, Stanton, Walker and Green (2006) have investigated 17 different SA measurement techniques which are classified as requirement analysis, freeze probe, real-time probe, self-rating, observer-rating, performance measures and process indices like eye tracker techniques. Also, these techniques are evaluated by validation studies. It is concluded that most of the measurement procedures lack construct validity[1] since they are not validated by further studies. Consequently, in this thesis two explicit measurement techniques tapping on different features of SA are used in the experiments, Situation Awareness Global Assessment Technique, SAGAT (Endsley, 1995a) and Situation-Present Assessment Method, SPAM (Durso, Hackworth, Truitt, Crutchfield and Manning, 1998). Due to problems of validity, two different techniques were used together as Çak (2011) suggested. SAGAT is selected due to its widespread application in the domain and SPAM is selected for its capability to capture the dynamics of the phenomena. Another reason for selecting these techniques is due to the fact that SAGAT is a

[1] Whether a scale measures or correlates with the theorized psychological scientific construct

freeze probe technique for obtaining Offline SA scores and SPAM is a real-time probe technique that is used to measure Online SA scores.

In the review mentioned above, Salmon et al. (2006) states that SAGAT is the only SA measurement technique with adequate validation studies. It is a freeze offline probe technique in which the task is interrupted and the operator is asked questions about the frozen environment. Obviously, this technique is applicable for simulator environments only. SAGAT queries involve questions to reveal accuracy in the three different levels of SA from Endsley (1995b)'s model. These levels are perception of the current situation, comprehension of it and projection of the future.

The other technique used in this thesis, Situation-Present Assessment Method (SPAM), is a real time probe technique in which the mission is not frozen during the SA queries. The operator is notified about the query during the task, but the query is not asked till the operator accepts. This method is less intrusive to the task performance and adds less workload on ongoing task. However, as one aim for the Çak's study was to find the correlates of SA in high workload condition, the option to wait till the workload gets low was eliminated. Participants were expected to answer as soon as possible. Consequently, SA queries were administrated under high workload conditions. SPAM technique is carried online such that the operators can find the answers for the queries while the task goes on. Durso et al. (1998) suggest that the operators do not necessarily need to keep all the information in the memory to attain SA but it is enough to know where to find that information. In SPAM, response times are obtained as well as the accuracy. Durso, Bleckley and Dattel (2006) looked for the incremental validity of SA on Air Traffic Controllers' performance and revealed that SPAM method contributed their statistical model's predictive power while off-line query models did not introduce that improvement. Moreover, latency in response turned out to be an important predictor of the performance, indicating underlying dynamic cognitive mechanisms. In this thesis work, a combination of SAGAT and SPAM methods as developed by Çak (2011) is used on a scenario adapted for the flight simulation environment. Further details on the modified scenario and SA queries will be given in the Appendix A.

2.2 Individual Differences

This work explores the relationship between individual differences and SA measurements. In this respect, working memory capacity, attention mechanisms - namely inhibition and divided attention- and expertise are investigated in terms of their definitions and connections to SA.

2.2.1 Working Memory Capacity

Working memory is the theoretical construct that can be broadly defined as storage of relevant information in the mind for a limited time in which complex functions on stored information can be carried out (Baddeley, 2000). Working memory is conceptualized as a broad term such that all complex cognitive activities requiring calculations use working memory. Some of these complex functions are reasoning, comprehension and learning. A prominent WM model is developed by Baddeley and Hitch (1974) as a three-component model taking the concept of short-term memory (Atkinson & Shiffrin, 1968) one step ahead. It contains a central executive module controlling two slave systems, namely the visuospatial sketchpad and the phonological loop. Slave systems handle the two modalities; vision and hearing where the central executive performs executive functions such as planning, problem solving, and verbal reasoning. Later, Baddeley revised his model as the multiple component model of working memory (Baddeley, 2000). The new component, episodic buffer, is introduced for associating information from different domains (vision, spatial and verbal) to form a portion of real life as a scene. The main construct of WM from Baddeley and Hitch is still influential in the current literature and conserves its validity throughout the years.

In applied cognition literature, most of the studies about Working Memory focus on Working Memory Capacity (WMC) due to the functional interest in working memory limitations. The famous study by Miller (1956) states "magic number 7 ± 2" chunks of information for WMC, whereas Cowan (2001) proposes that the magic number is less than 7 and people can only handle approximately four items at a time. Assessment of individual differences in working memory capacity is generally done by measuring the number of information (words, letters, numbers, etc.) that can be simultaneously stored and processed by the individual. In the literature, there are a limited number of WM span tasks that have been widely used as a successful predictor of both higher order and lower order cognitive task performance. Except for the tasks developed for working memory deficits, reading, counting and operation spans are the most commonly used tasks (Unsworth, Heitz, Schrock, Engle, 2005). In order to investigate the different working memory tests, Conway, Kane, Bunting, Hambrick, Wilhelm and Engle (2005) review counting span, operation span, and reading span tests and compare them within themselves and with Raven's Advance Progressive Matrices. In the reading span task (Daneman & Carpenter, 1980), a sequence of sentences is presented and last words from each sentence are expected to be recalled at the end. Similarly in the counting span task (Case, Kurland, & Goldberg, 1982), participants are requested to report the number of objects successively presented. Lastly, operation span task (Turner & Engle, 1989) contains sequential tasks of solving equations and remembering words in the order they were presented. Conway et al. (2005) concludes that counting, reading and operational span tasks are reliable and valid measures of WM span. Unsworth et al. (2005) forms another kind of WM span task, automated version of the operational span task (AOSPAN). In this version of operation span, the participant carries out the task on a mouse driven computer. Automated version of this task will be described in detail.

For the AOSPAN task, three practice sessions take place in which letter span task, a simple mathematical equation solving task (e.g., (2*2) – 1 =?) and these two tasks together are carried out. After participants completes the practice sessions, a total of 75 letters and 75 math problems within set sizes ranging from 3 to 7 are presented as the real task. The participants have to attain a 85% accuracy in math problems at all times. AOSPAN score is calculated as the sum of all perfectly recalled sets. At the end of the task, total number correct answers, math errors, speed errors, and accuracy errors are displayed.[2] Unsworth et al. (2005) concluded that AOSPAN highly correlates with other WM span tasks and has both good internal consistency[3] and

[2] This test is translated to turkish in the scope of Çak's studies.

[3] Whether several items that propose to measure the same general construct produce similar scores

test-retest reliability[4]. Consequently, AOSPAN was introduced as a reliable indicator of a broader construct of WM among several WMC tasks. All WMC tasks have specific features that affect the obtained results. For the investigation of SA as a complex cognitive concept AOSPAN task was more suitable compared to above listed tasks in the literature, because it is a complex task and it shows differences among simulator pilots. This is enabled by the composition of mathematical equation solving and letter span tasks together. Moreover, AOSPAN task is carried on a computer so that it is administrated and scored easily. Response times can also be obtained using this automated version. AOSPAN task is used for measurement of WM in this study due to the findings summarized above.

As given above, working memory capacity is regarded as a core predictor of several cognitive abilities like executive attention theory (Engle & Kane, 2003). However, Kane, Poole, Tuholski and Engle (2006)'s study on different search strategies reveals that WM capacity was correlated with not all but a limited dimensions of executive attention. The executive attention processes related with WM can be given as (1) attending stimuli and goal outside of conscious focus under interference from prior experiences or habit, (2) handling response competition and controlling context-inappropriate responses, and (3) maintaining conscious focus during distraction (Kane et al., 2006). Given classification supports the hypothesis of importance of WM for SA since all listed executive attention processes are crucial for SA. As mentioned above, AOSPAN test will be used to capture WM as an underlying factor of SA.

Besides the expected relationship between SA and working memory capacity, attention mechanisms are also predicted to be important for a better SA. Simulator pilots, like professional pilots, observe several equipments and weather conditions together and carry out multiple tasks in the cockpit. Consequently, inhibition and divided attention capacities are also possible abilities that underlie a successful SA.

[4] Yielding same results in further tests under same conditions

2.2.2 Attention Capacity

Attention can be defined as the general ability to attend to activities happening in the outside world or within the mental environment. The crucial task for attention is the successful allocation of limited mental resources within task requirements to handle vast amount of data in those environments. For the complex process of SA, importance of attention is stated in several works (Sarter & Woods, 1991; Endsley, 1995a).

In order to cover the broad concept of attention, two sub-mechanisms are suggested to define it more precisely, namely selective and divided attention (Cherry, 1953). It is proposed that picking the relevant information from a whole body of information and filtering the unnecessary ones is done by selective attention. Divided attention, on the other hand, is defined as the ability of distributing the limited capacity of attentional resources over more than one task. Both of these capabilities are very important for the processes of SA in the flight simulation environment where pilots observe several conditions about the flight. Especially, observing both different flight equipments and the environment outside of the aircraft requires a well managed divided attention capacity.

Theories of attention are studied beginning from James (1890). According to James, attention is briefly described as "withdrawal from some things in order to deal effectively with others". In the psychology literature, one of the important pieces of research about attention is the information theoretical approach introducing a limited capacity channel by Broadbent (1958), also known as the Filter Theory. This filter only passes the relevant information and keeps the irrelevant ones out. It is later improved by Treisman (1960) such that irrelevant information is just suppressed but not totally ignored. These theories of attention are known as early selection theories. There are late selection theories, on the other hand, where attention filters the information after semantic pre-processing (Deutsch & Deutsch, 1963). Later, a flexible bottleneck theory of attention is proposed by Johnston and Heinz (1978) where the bottleneck can be actively switched for early or late selection, combining both of these approaches. Attention control mechanisms are introduced to enable this switching between phases of selection. Similarly, Lavie, Hirst, de Fockert, and Viding (2004) distinguishes high and low perceptual load conditions and propose a dual filter for each phase of selection under different perceptual loads.

In the literature, spotlight and zoom lens metaphors are also influential. LaBerge & Brown (1989) describe attention as a spotlight which points to a piece of information at a time, demonstrating the selective behavior. The zoom lens metaphor, on the other hand, could be regarded as an improvement over the spotlight metaphor. Eriksen and St. James (1986) use the zoom lens metaphor to explain the properties of attention. When focused on something, more details can be accessed on a small portion of the whole information space. However, when not focused, a larger area of space can be accessed with lower details.

Apart from the studies exploring the nature of attention, effects of expertise are also introduced in attention studies leading to a distinction of automatic and controlled attention management. Novice performance is governed by controlled attention where careful and effortful performance is needed. This controlled attention is limited by capacity. However, expert performance is governed by automatic attention and relatively less restricted compared to controlled attention. Familiarization with tasks increases the capacity limits for attention. (Posner & Snyder, 1975; Schneider & Shiffrin, 1977). In this study, combined effects of expertise and divided attention capacity are investigated as potential indicators of SA performance.

The concept of individual differences in attention capacity is also investigated in several studies, especially in applied settings. Gopher (1982) shows that participants who are more successful in a selective attention dichotic listening task were also more successful in flight training. Jones and Martin (2003) find that computer users with lower ability of divided attention are more likely to forget to save their work on the computer. These studies about expertise and individual differences provide strong evidence for attentional mechanisms underlying SA. In this study, two different tasks are used to capture individual differences in attention; Stroop task for prepotent response inhibition and Çoşkunöz visual attention test (Er, Sümer, Koku, Mısırlısoy, Coşkan, Erol-Korkmaz, Sümer, Ayvaşık, Eriş, 2011) for divided visual attention.

Simulator pilot's SA performance is expected to be connected to several attentional mechanisms. Stroop task (MacLeod, 1991) is one of the tasks used to assess prepotent response inhibition which is a crucial feature of executive functioning for attention (Friedman & Miyake, 2004). As noted in Çak (2011), pilots need inhibitory mechanisms during their performance in flight simulation. Every part of the flight requires different procedures and pilots have to focus on different sets of equipments during these parts. In ILS approach[5] part, for example, simulator pilots have to carefully observe glideslope and localizer information[6] instead of checking heading information. As in this case, ignoring the commonly used equipments and focusing on a specific set of equipments is an exemplar case of prepotent response inhibition.

Stroop task is used to determine the Stroop effect (MacLeod, 1991) which is the delay due to prepotent response inhibition. In the task procedures, participants are asked to name the colors of strings. It is observed that participants with lower inhibition abilities have longer delay in the incongruent cases when they have to name the color of a string which is another color name. In this case, prepotent response of word reading has to be suppressed in favor of the task goal, color naming. Delay in the incongruent cases compared to congruent cases (same color name as strings) or neutral cases (symbols as strings) is evaluated as a measure of inhibition capability in the context of attention.

Another attentional mechanism tested in the scope of this thesis is divided visual attention. Simulator pilots, unlike professional ones, do not effectively use radio due to the environmental and procedural differences. However, similar to professional pilots they have to visually observe several conditions inside and outside of the plane. Weather conditions, several equipment displays and flight charts are to be simultaneously followed. Eriksen and Yeh (1985) propose a theory for allocation of attention in visual field. According to their study, visual attention can cover a larger area of interest with low resolution. When focused on small portions of the visual

[5] Approach using vertical and horizontal guidance from the airport transmitter

[6] Vertical and horizontal guidance showing the deviations from the predefined approach line

field, resolution increases for areas of attention. Simulator pilots follow different flight equipments and also the weather conditions at the same time. There are cases of loss of SA when pilots focus on specific equipments and fail to follow some others (Çak, 2011). Participants with high SA performance are expected to have higher divided attention visual capabilities. In an effort to test this relationship, Çoşkunöz visual attention test developed by Er, Sümer, Koku, Mısırlısoy, Coşkan, Erol-Korkmaz, Sümer, Ayvaşık & Eriş (2011) is used. In this test, participants follow and respond to two different tasks unfolding simultaneously on the left and right sides of the screen. These tasks are not synchronized and have to be attended together. Further details of the Coşkunöz visual attention test can be found in Chapter 3.

In this study, Stroop task and Çoşkunöz visual attention test are used to assess inhibition and divided attention capabilities of the simulator pilots. These abilities are possible candidates of attentional mechanisms underlying SA. Simulator pilots with higher inhibition and divided attention capabilities are expected to be more successful in SA. Similarly, expertise is also expected to be an important factor in good SA. Expertise provides automaticity for the management of several attention-bearing tasks, and hence simulator pilots with higher flight hours are expected to benefit more from such automated skills.

2.2.3 Expertise

Expertise, as in any task, is a significant factor for success in piloting. In this respect, ranking systems for pilots in airlines are based on years of piloting and hours of flight. Experienced pilots working for airlines are expected to take more responsibilities and give vital decisions during the flights. Simulator pilots, similar to professionals, have ranking systems in virtual airline communities. As they have more flight hours, they get higher rankings; however, the responsibilities that they take do not increase accordingly.

Except for the individual cognitive differences given above, effect of expertise is an important aspect of complex-task performance. In their reviews, Ericsson and Charness (1994), and Ericsson and Lehmann (1996) stated that regular careful practice performed for long enough provides expert-level performance in a domain. Based on this background, acquiring domain-specific skills for efficiently encoding the presented information to the long-term memory was identified as a core factor of expertise. In terms of SA, Endsley (2006) points to the significance of expertise by denoting mental models that develop in a pilot's mind. Pilots have an increasing picture of whole details about the flight as they have more experience and it gets easier to understand the current situation, carry out the tasks, and predict the future. Accordingly, in this study participant's expertise is evaluated in terms hours of flight and regarded as a predictor of SA performance.

2.3 Individual differences in SA and Underlying Factors

In order to learn more about the phenomena of Situation Awareness, exploring the individual cognitive differences is a necessity. The literature reviewed above does not converge on a clear conceptualization of SA. For this reason, interactions of SA with working memory capacity, attention mechanisms and expertise as individual differences are used to investigate SA in cognitive terms. Below is a summary of SA-related work tapping on these individual differences.

Endsley and Bolstad (1994) conduct a study on experienced fighter pilots to compare and associate their several attributes (categorized by spatial, attention, memory, perception, and cognitive) to SAGAT scores. They conclude that spatial and perceptual skills are important for SA. However, they fail to find any relation between SA and memory or analytical skills. An important study that explicitly focuses on the relationship of WM and SA is Johannsdottir (2004). In this study, the hypothesis of active role of working memory in maintaining SA is supported with experimental results. Central executive control, the verbal and spatial WM systems are found to be involved in maintaining SA. The effect of expertise on SA is not investigated in Johannsdottir's (2004) study. In addition, Johannsdottir carried SA measurements in simplified tasks, which do not have any operational basis. Participants for Johannsdottir's study were also not a SA-related group but university students. Another relevant study is by Sohn and Doane (2004), in which the authors explore memory processes of flight SA for examining the role of WM capacity, Long-term Working Memory (LT-WM), and expertise. The concept of LT-WM is expressed as a structure functioning like a WM but storing considerable amount of information related to learned skills (Ericsson & Delaney, 1999; Ericsson & Kintsch, 1995). Capacity of WM was increased by extensive utilization of long time memory through expertise. Based on empirical data from Sohn & Doane's study, it is concluded that WM capacity and LT-WM affect SA performance; however their effects change with expertise. For novices with 85,7 hours of average flight time, spatial working memory is the most predictive feature, whereas for experts with 1116,8 hours of average flight time, spatial long-term working memory component turns out to be more predictive. However, the construct of LT-WM is theoretically weak due to its domain-specific definition. It is measured in connection with abilities from piloting and do not represent a general ability. Therefore, this construct is not used in the current study.

Processes underlying SA is also investigated in the driving domain. Gugerty (2011) states that focal processes underlying SA such as attention allocation, event and risk comprehension and task management improves with training and expertise. For the ambient processes, he suggests that the ability of detecting sudden peripheral events may be related to SA. Gugerty's study reviews studies in the literature that aim to find the components of SA, but does not point to cognitive constructs like working memory.

A recent study on this area has been conducted by Serkan Çak (2011). Çak's study pursues the relationship between SA and individual cognitive differences. In this study, 36 professional pilots were subjected to SA and cognitive capacity tests. For SA measurements, participants performed a flight including novel events and special conditions on a full flight simulator. Çak's study was used as a starting point for the current study. The flight scenario, SA queries and cognitive capacity tests were used for simulator pilots as similar as possible to be able to have comparable results to professional pilots. Further details of the flight scenario developed by Çak and the adaptation applied for the purposes of this thesis can be found in Appendix A.

In Çak's study, SA measurements were administrated using offline and online SA queries. These queries were carried out during the flight. In order to assess individual cognitive differences three different tests were administrated; AOSPAN, Stroop and a pilot specific Choice Reaction Time Task with Dichotic Listening (Çak, 2011). It has been found that working memory capacity and expertise account for 58% of variability in offline SA scores (Adjusted R^2 = .58 (F (4, 35) = 12.81, p = .00). Inhibition and divided attention capacities did not show any predictive power for offline SA scores. However, working memory capacity was the most successful predictor (β = .675, t(31)= 5.31, p<.00). Expertise was the other main predictor (β= .278, t(31)= 2.35, p<.05).

Another measurement for SA was done by using online queries. For online RT measurements, 52% of variability is explained by inhibition, divided attention and expertise. (Adjusted R^2 = .52 (F (4, 35) = 10,29, p = .00). Predictors for online RT was given as expertise (β = -.470, t(31)= -3.73, p<.001), divided attention (β = .313, t(31)= 2.25, p<.05) and inhibition (β = .260, t(31)= 2.058, p<.05). No significant correlation was found between the predictors (individual cognitive differences and expertise) in the regression models. Hypothesis of Çak's study were supported by the experiments. Working memory capacity, inhibition, divided attention and expertise have predictive power over SA measured by online and offline queries. In short, for the offline SA scores, working memory capacity and expertise are the strongest predictors, whereas for the online SA scores, inhibition, divided attention and expertise are the strongest predictors.

Apart from the behavioral experiments, Çak's (2011) research also includes an eye tracking study to investigate the effects of individual differences. Several hypotheses were tested using fixation count and fixation duration measures obtained during four different tasks carried out by ten expert pilots in two categories of expertise. It was aimed to find difference between pilots with respect to gaze patterns between high or low expertise, high and low working memory capacity, long and short inhibition delay and lastly high and low divided attention capacity in terms of fixation count and fixation durations. However, only one of the hypotheses, the difference between more expert and less expert pilots in terms of fixation duration is supported. More expert pilots fixed their sights for a shorter duration as compared to novices, which indicates their proficiency on detecting what they need. Effects of working memory capacity, inhibition delay and divided attention capability were not observed in the eye tracking experiments.

Cognitive explanation of the phenomenon of SA was not explored enough by experimental methods in the previous studies. Only a limited portion of previous studies focused on experimental investigation of the cognitive aspects of SA; however those studies were not conducted with real operators in real task environments. This gap was aimed to be filled by Çak's doctoral studies. Çak's research was conducted with professional pilots on full flight simulator systems. Inspired by Çak's work, this study investigates if similar findings apply to simulator pilots in a desktop flight simulation environment. Using the empirical background set by the literature reviewed above, the present study explores the dimensions of individual differences and expertise in relation to SA in real task environment, and looks for significant relations among them. To our best knowledge, simulator piloting was not studied before in the exploration of SA. This study investigates the connections between professional and simulator piloting as well as contributing the cognitive exploration of SA.

2.4 Optical Brain Imaging (fNIR) for Workload Assessment

Near infrared spectroscopy is an optical method that can non-invasively monitor cerebral oxygenation changes. Its application to monitoring brain metabolism was first described by Jobsis (1977). Later, functional near-infrared (fNIR) spectroscopy has been developed in order to investigate brain activities (Chance, Zhuang, Alter, Lipton, 1993; Villringer, Planck, Hock, Schleinkofer, Dirnagl, 1993; Villringer and Chance, 1997). Changes in the concentration of deoxygenated and oxygenated hemoglobin are monitored with the change of light absorption property of the tissue. Hemoglobin, unlike most of the biological tissues, absorbs light waves in the near infrared range and this makes optical observation of the change in the concentration of hemoglobin possible (Izzetoglu et al., 2005). In connection with changes in hemoglobin concentration, tissue oxygenation and metabolism can be investigated during cognitive tasks (Villringer & Chance, 1997). Cognitive tasks such as attention, working memory, target categorization, and problem solving can effectively be monitored by the fNIR technique. Findings from fNIR studies are also compatible with the blood oxygenation level dependent (BOLD) signal of fMRI (Cooke, Pringle, Pederson, Connor, 2006; Bethke, Valenti, How, 2007; Tvaryanas, Thompson, 2008). An obvious advantage of this technique is that brain activations can be obtained in naturalistic task environments, especially for the special environment necessary for the simulated flight task.

The current literature shows that fNIR is an effective technique to assess mental workload in naturalistic task environments. Sassaroli, Zheng, Hirshfield. Girouard, Solovey, and Jacob (2010) carried out an fNIR study where participants were asked to count the number of different colors on a rotating cube. The image of a cube was showed such that the top face was always visible, and as it rotates the side faces are shown one by one. The bottom face was hidden. Participants had to remember the number of each different color. This task was designed in such a way that it contained four different difficulty levels as the number of different colors increased. After the analysis of the fNIR signals under different mental workload levels, significant differences in brain activations in terms of oxygenation were observed. Sassaroli et al. have reported that fNIR has the potential to characterize various levels of mental workload.

The use of operators' self-rating of their workload during or after completing a task is often subjected to criticism for not providing an objective measure of mental workload. This motivated Menda, Hing, Ayaz, Shewokis, Izzetoglu, Onaral, and Oh's study (2010) to employ fNIR technology in an effort to devise a more objective measure of mental workload by comparing brain activities induced by different control interfaces. While participants were conducting tasks in a realistic environment, their brain activities at the dorsolateral prefrontal cortex (DLPFC) were collected non-invasively. Two different interfaces for controlling a near-earth UAV system were compared in terms of overall workload ratings, mental demand ratings and oxygenation changes. Although no difference was observed in subjective tests (i.e. overall workload and mental demand), the fNIR analysis found a significant difference in average oxygenation changes in favor of the interface where the UAV was controlled through a view augmented with a virtual model of the environment. In the analysis, fNIR data obtained from voxel 4 was found to be particularly responsive to this difference, which is a region associated with concentration, attention and working memory (Izzetoglu et al., 2007; Ayaz et al., 2009).

An aircraft identification task was also used to investigate the relationship of the hemodynamic response in dorsolateral prefrontal cortex with mental workload (Bunce, Izzetoglu, Ayaz, Shewokis, Izzetoglu, Pourrezaei & Onaral, 2011). In the study, participants had to determine friendly or hostile aircrafts in a Warship Commander Task with increasing difficulty as more aircrafts are needed to be identified. A significant increase in fNIR oxygenation measurements was observed as the mental workload increased with task difficulty. In a similar study with Air Traffic Controllers (ATCs), mental workload assessment was done again by the fNIR technique (Ayaz, Shewokis, Bunce, Izzetoglu, Willems, Onaral, 2012). ATCs participated in standardized (n-back) and complex cognitive task (air traffic control). It is reported that with the increase in task difficulty, both for standardized and complex cognitive task, increase in perceived workload can be observed significantly in the fNIR signal. For the n-back task, oxygenation readings from voxel 2 showed reliable increase with n-back condition. This voxel is associated with AF7 in the International 10–20 System, which is located within the left PFC (inferior frontal gyrus). For the ATC task, participants were controlling 6, 12 and 18 aircrafts during the increasingly difficult task. Main effect of task difficulty was observed in voxel 8 located above the medial PFC, which is often associated with decision-making processes (Miller, Freedman, & Wallis, 2002). Statistical analysis results showed that both perceived workload and fNIR signal significantly differ as the task got more difficult.

These studies suggest that fNIR optical brain imaging technique can be potentially used to monitor changes in mental workload as a function of changes in cerebral oxygenation at the prefrontal cortex. In this study, the fNIR technique was used to investigate if a greater level of oxygenation was induced by the high versus low workload episodes of the simulated flight scenario.

CHAPTER 3

METHOD

In this chapter, the experimental setup used in the study will be described. It consists of a simulated flight task, AOSPAN, Stroop and Coşkunöz visual attention tests for measurement of SA, working memory capacity, and inhibition and divided attention capacities, respectively. An optical brain imaging study with fNIR technique that is carried during above mentioned tasks with a limited number of participants will also be described.

A simulated flight, AOSPAN, Stroop and Coşkunöz visual attention test tasks were carried out as behavioral tasks. Similar to the study by Çak (2011), cognitive understanding of the phenomena of SA was explored by analyzing SA measurements together with working memory capacity and attention mechanisms. In order to have comparable results between professional and simulator pilots, this study was based on Çak's doctoral studies (2011). Similar experiments, except for divided attention, with minor modifications were used. Multimodal divided attention task was changed with a unimodal one because in Çak's study a group of professional pilots were the participants who have to pay attention to both visual and auditory modalities during flight. However, the participants in this study, simulator pilots do not use radios extensively during flight and they don't pay any attention to auditory modality. For this reason, a different divided attention task, Coşkunöz visual attention test was introduced. The first task is was simulated flight for SA measurement, the second is Automated Operational Span Task (AOSPAN) for working memory capacity, and the last two are Stroop task and Coşkunöz visual attention test for attention mechanisms measurements.

For assessment of the level of neural activity at simulator pilots' dorsolateral prefrontal cortex, an optical brain imaging study with fNIR has been performed during SA and cognitive capacity tasks with a limited number of participants. Since these tasks are used for fNIR measurements, no other tasks are designed for this additional study.

3.1 Preceding Study and Modifications

In Çak's doctoral studies, SA was investigated from a cognitive perspective with professional pilots. This study was built on that work, but this time with simulator pilots. Çak conducted SA measurements with a flight scenario and three cognitive capacity tests which are AOSPAN for Working Memory Capacity (WMC), Stroop Task and a pilot-specific Choice Reaction Time with Dichotic Listening for attention mechanisms (Çak, 2011). In order to have comparable results from professional and simulator pilots, experiments in this study was administrated as similar as possible to Çak's. The three of the experiments, SA measurement, AOSPAN and Stroop Task were replicated with little or no modifications. For divided attention measurement, a new experiment Coşkunöz visual attention test based only on visual modality is carried out. Further details of the tasks will be given in the following subchapters.

The flight scenario for SA measurement in his studies was designed to be cognitively compelling such that some novel events were introduced during the flight. Since the original flight scenario was designed for professional pilots, it is not used as it was. Expert simulator pilots worked as Subject Matter Experts (SME) in order to review and modify some parts of the scenario for simulator pilots. For this purpose, several flights have been conducted with SMEs to validate the scenario for a simulation environment. Consequently, minor changes were applied to adapt the scenario. General information about the scenario will be given in the related subchapter. Detailed information and modifications on the original scenario can be found in Appendix A.

The original scenario was administered by instructor pilots in Çak's work. In the present study, however, the administration is done by the researcher. For this reason, instructions for the instructor have been extended with SMEs in order to assure replicability of the experiment.

AOSPAN and Stroop Task as widely used tests in several studies have been used with no modifications. The Choice Reaction Time with Dichotic Listening test was not reused since it is not a practice that simulator pilots perform. A different experiment, where attention was divided only in visual modality, was used.

3.2 Pilot Study

Pilot study has been carried out for two purposes in this study. First, necessary modifications for the simulated flight task have been determined by three flights carried out with SMEs. SMEs are selected among the expert simulator pilots in Ankara who are proficient about the simulator piloting. SMEs participated in these flights of the original scenario to ascertain feedback for modifications. These modifications are listed in Appendix A. After that, the scenario was modified with respect to the differences in practices between the simulator and professional piloting. Second, other two complete flights of the modified scenario together with AOSPAN, Stroop and Coşkunöz visual attention test were carried out with ordinary simulator pilots. It enabled us to go over the whole tasks for behavioral measurements.

The last two sessions of pilot experiments contained all tasks performed by simulator pilots to go over the whole procedures. During the last session brain imaging data was also collected with fNIR technique. Obtained data was analyzed and found to be useful for our purposes. Further details of the data analysis will be given in the following chapter.

3.3 Participants

For the behavioral experiments, thirty five simulator pilots (all male) with a moderate to an advanced experience level participated. They were selected through online simulation communities. Simulator pilots who also have professional piloting experiences were not accepted in the study[7]. Participants received a little amount of money for their participation. To check their proficiency, all participants were asked to complete a pre-flight scenario at home and submit the flight logs. This pre-flight included similar situations critical to the scenario of the experiment. Details of the pre-flight scenario can be found in Appendix A. All participants were native speakers of Turkish. Their mean age was 30.7 and their average total flight hours was 1356[8]. Biographical information about the participants is given in Appendix B. To have a comparative look at Çak's study, information about participants in his study is also given in Appendix B. These data are given so that results from Çak's study can be comparatively evaluated with respect to this thesis in terms of professional and simulator pilot's expertise and cognition related conditions.

fNIR brain imaging study has been conducted with ten participants at the Turkish Armed Forces Modeling and Simulation R&D Center at METU. Since the fNIR device was located in Ankara, only the participants from this city were included in the study.

3.4 Apparatus

The SA measurement is conducted with a flight scenario in Microsoft Flight Simulator 2004. The flight scenario adapted from Çak (2011) requires real-time manipulation of weather and aircraft conditions by the instructor. In order to find a suitable platform that enables online change of varying parameters, different

[7] One participant working as a professional Air Traffic Controller and two participants who were student pilots participated in the study. Data from these participants were not found to be outliers.

[8] Total flight hours for simulator pilots include piloting experiences from different platforms.

simulation platforms and extensions have been investigated. Microsoft's Flight Simulator 2004 which has an instructor panel for real-time manipulation is chosen.

The experiment setup was designed to be basic to minimize several participants' familiarization problems. The simulation computer was a standard laptop computer with 15.6" screen and high graphics performance. A standard joystick, of a type which is generally used by moderate level simulator pilots, was attached to the simulation computer. For manipulation of the weather and aircraft conditions, another laptop computer was connected to the simulation computer and used as an instructor station by the experimenter. Cockpit setup on simulation computer and instructor panel on experimenter's computer can be seen in Figure 3.1 and Figure 3.2 respectively.

The communication between the participant and the experimenter was recorded during the flight in order to acquire response time data for online SA queries administered.

Figure 3.1 A sample screenshot showing cockpit setup on simulation computer

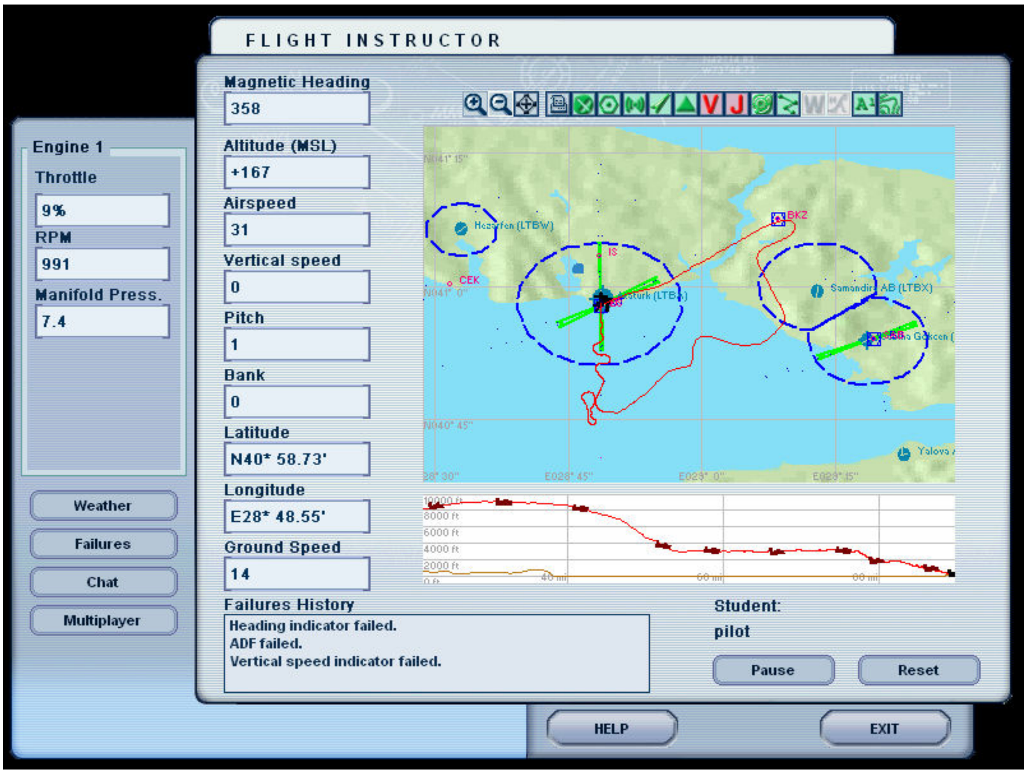


Figure 3.2 A sample screenshot showing instructor panel on experimenter's computer

The communication between the participant and the experimenter was recorded during the flight in order to acquire response time data for online SA queries administered.

Working memory capacity and attention mechanisms measurement tests were also administrated on the simulation computer described in the previous paragraph.

The behavioral experiment was conducted in four different cities in order to obtain adequate participation. These cities are Ankara, Konya, İstanbul, and Antalya. Numbers of participants from each city are 11, 3, 8, and 13, respectively. All participants in different cities took the experiment in a room alone with the experimenter under similar lighting conditions. Experiment setup was also carried to these locations and experiments were administered with the same setup in order to assure ecological validity.

For the brain imaging study, the prefrontal cortex of the participants were monitored by using a continuous wave fNIR system called fNIRS Imager 1000, which was first described by Chance et al. (1998), further developed at Drexel University, and manufactured by fNIR Devices LLC (Potomac, MD; www.fnirdevices.com). A flexible sensor pad, given in Figure 3.3, containing 4 light-emitting diodes (LEDs) and 16 sensors, a control box for hardware management and a computer for data acquisition are used in the system. Dorsal and inferior frontal cortical areas underlying the forehead are measured for brain activity in 16 locations by the arrangement of the photodetectors and the light sources in the sensor pad. COBI Studio Software (Ayaz, Shewokis, Curtin, Izzetoglu, Izzetoglu, Onaral, 2011) was used for fNIR data acquisition.

Figure 3.3: The fNIRS System sensor pad

On the sensor pad, 4 light sources (LEDs) emits light in near-infrared wavelengths (650-900 nm). Four photodetectors around each LED measures the reflected light by the tissue under them. The infrared light is mainly absorbed by hemoglobin structure but not the skin, skull, cerebrospinal fluid and brain tissue (Obrig & Villringer, 2003). Consequently, hemodynamic changes due to the changes in the relative concentrations of oxy- and deoxy-hemoglobin can be observed by the sensor above the specific brain area.

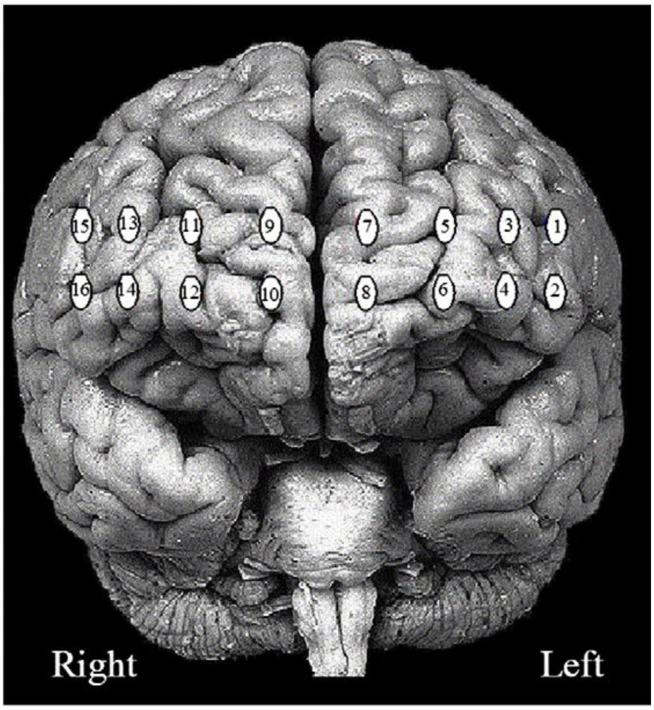


Figure 3.4: 16 measurement locations registered on the surface of the brain adopted from Ayaz et al. (2011, p.39)

For the identification of the fNIR data, several markers are used to assess the different segments of the behavioral experiments. Blocks have been formed with respect to markers that correspond to specific segments of the flight and further analysis has been carried on these blocks of data. For each participant, the 16 channel raw fNIR data were passed from a finite impulse (FIR) low pass filter with order 20 and a cut-off frequency of 0.14 Hz. This filtering attenuates the effects of high frequency noise and respiration/cardiac cycle effects. Saturated channels (if any) where light intensity at the detector was higher than the analog-to-digital converter limit were excluded. For filtering out the noise due to head motion, sliding window motion artifact filter (Ayaz, Izzetoglu, Shewokis, Onaral, 2010) was applied next. Lastly, in order to calculate oxygenation values (the difference between oxy- and deoxy-hemoglobin concentrations) modified Beer Lambert Law (Chance et al., 1998) is used. Oxygenation values are computed with respect to baseline measures obtained at the beginning of the flight while the participants were resting with their eyes closed. These analyses have been performed with the "fnirSoft Software" (Ayaz, 2010). Brain activation in the specific area can be inferred by the increase in oxyhemoglobin concentration with respect to de-oxyhemoglobin concentration, due to increasing oxygen demand. (Izzetoglu et al., 2005).

3.5 Design

In this section, the design of the four different task used will be described. These tasks are Simulated Flight, AOSPAN, Stroop, and Coşkunöz visual attention test in order. fNIR study, on the other hand, was aimed to observe brain activations during the above listed SA and cognitive capacity tasks. As suggested by Izzetoglu (2007), fNIR technology is an efficient method for assessment of cognitive activities like attention, working memory and problem solving. Thanks to the non-invasive structure of the device, participants' hemodynamic changes that occur in their brains have been recorded under the experimental conditions for SA, working memory capacity and attention mechanisms measurements.

3.5.1 The Simulated Flight

In the first task, participants were asked perform a simulated flight as a pilot for Cessna-172 fixed wing aircraft in flight simulation environment. The modified scenario adapted from Çak (2011) was used. During the flight several uncommon events occurred increasing the mental workload. Questions about the current status of flight are asked for an assessment of the participant's situation awareness (SA). In the following paragraphs, Çak's scenario, and modifications thereof, will be briefly described and elaborated on in terms of the scenario duration, workload, novel events, duplicability and ecological validity.

Similar to the original scenario from Çak (2011), participants performed a medium length scenario. The duration of the flight simulations in general practice is 1 to 4 hours. According to Çak's study, the minimum length of a cognitively demanding scenario is roughly 75 minutes. Thus, the length of the simulation was planned in accordance to the Çak's study.

The adapted scenario from Çak (2011) comes with novel events, which increases the workload compared to a typical flight. The take-off and climb phases are designed to be standard and possess a low workload demand for quick familiarization at the beginning of the experiment. However, after the first 25 minutes, the weather becomes worse and several equipment failures should result in an increased workload for cruise, descent and approach phases of the flight. As Çak (2011) suggested, SA queries were administered during the high workload phases of the flight. Two different techniques were applied for SA measurement, online (SPAM) and offline (SAGAT). The original SPAM technique was modified by removing the "reject to answer" option in order to assure high workload during the SA queries (Çak, 2011). The original scenario was retained in terms of novel events experienced during the flight. The novel events were icing, rain, turbulence, crosswind[9], low

[9] A wind passing through the right or left side of the aircraft

visibility, low ceiling[10] and planned failures in the equipment (VSI[11], RMI/HSI Compass[12], and ASI[13]). The motive was to introduce high workload and stress to the flight, which reveals cognitive differences among participants.

The participants had to fly the aircraft under icy, windy, and low-visibility weather conditions (Çak, 2011). Due to icing above a specific altitude, the simulator pilots had to cancel the flight and turn back to the departure airport. Clouding conditions resulted in low visibility, which requires more effective use of flight instruments by the participants. Winds, on the other hand, affected the aircraft with turbulence and cross-wind. Both of these factors increased the effort required to pilot and the scenario complexity. Lastly, a low ceiling was an important challenge involved in landing, since participants needed to have visual contact with the runway at a given altitude. Under these conditions, participants had to execute a missed approach procedure[14].

Instrument failures were the other kind of novel events added to the scenario (Çak, 2011). Heading Indicator[15], ADF[16], VSI and ASI were set to fail in a fixed, predefined order. Participants were forced to use a magnetic compass in the cockpit in case of heading indicator failure, which is not a common practice. Failure of the ADF did not critically affect the flight; however participants had to follow the magnetic compass, the only navigational aid left. VSI and Airspeed Indicators are

---

[10] The height above the ground of the base of the lowest layer of clouds

[11] Vertical Speed Indicator shows the vertical speed of the aircraft

[12] Radio Magnetic Indicator/Horizontal Situation Indicator show the horizontal situation of the aircraft with respect to compass and radio transmitters on the ground

[13] Air Speed Indicator shows the speed of the aircraft with respect to surrounding air.

[14] Flight procedures prescribed when an aircraft fails to land after completing an approach

[15] Equipment showing the direction that the aircraft's nose is pointing

[16] Automatic Direction Finder shows the aircraft's relative direction to a radio transmitter

important gauges of aircraft motion, the failure of which forced participants to make extra calculations with regards to the time to explicitly find these values.

The original scenario included a special case, where participants were required to fly to an intersection point, with no explicit navigational markers (Çak, 2011). This special point was at a specific distance from a radio station and the participants were expected to track a special route. However, this is again not a common practice for both professional and simulator pilots and they spent a lot of mental effort to find a navigation solution. If they had failed to find a solution in 10 minutes, the experimenter helped them by providing the necessary heading information.

SA measurement over a scenario inherently has dynamic characteristics. Even if participants do the same actions, due to their timing and all other flight conditions at that time, each flight happens to be different. In order to get reliable SA measurements, test instructions have been very clearly defined in connection to specific events during the flight. Since this study was administered by the researcher, Çak's (2011) instructions for instructor pilots were carefully evaluated and elaborated for replicability. Details of the modifications of the original scenario can be found in Appendix A.

During the flights, the researcher fulfilled several roles, as experimenter, flight instructor, and Air Traffic Controller. Test instructions for the experimenter can be found in Appendix C. This situation has been observed and approved by SMEs. As declared previously, the scenario used, was adapted from Çak (2011) with minor modifications. In this and the following two paragraphs important aspects of SA measurement will be presented. The scenario consisted of a typical flight between two airports. However, at a given point due to icing conditions, pilots had to terminate the flight since the aircraft is not equipped to fly in these conditions.

Consequently, the participants were instructed to fly to a holding-point[17] and to land back at the take-off airport. This event separated the flight into low and high workload segments and SA measurements were carried out in the latter segment. The holding point was an intersection point at which participants had to develop his own strategy to fly to using several instruments. After reaching and having a hold at this point, instruments providing directional information were frozen and the participants were told to approach for landing. Afterwards, due to the low ceiling, the first landing could not be achieved, and missed approach procedures had to be executed. At the second try for landing, airspeed and vertical speed indicators became inoperative while weather conditions allowed for landing. An overview of the flight route can be seen in Figure 3.3.

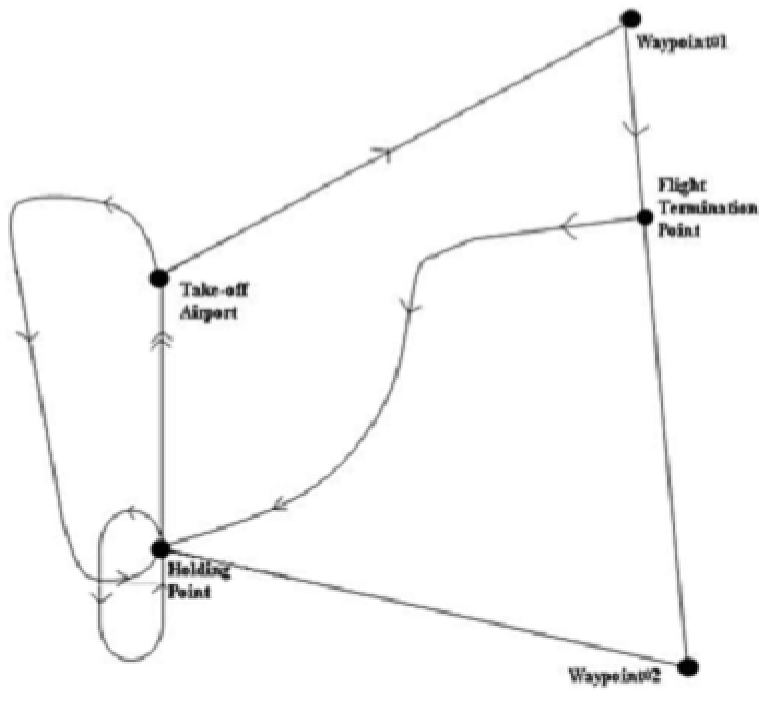


Figure 3.5: Flight Route in the Scenario adopted from Çak (2011, p.50)

The assessment of SA in the experiments was done by a combination of SAGAT and SPAM techniques (Çak, 2011). The content of the questions and their timing were

[17] A predefined point for stationing aircraft in flight

prepared by Çak together with SMEs in order to be able to extract the important aspects of the flight. Queries were administrated in two ways, online and offline. Online queries were motivated by the SPAM technique, where the participants were asked to respond as the simulation runs. The response time and accuracy data were gathered. Offline queries, on the other hand, were motivated by SAGAT procedures, in which the simulator was frozen and the participants were asked to fill out a questionnaire about the flight only concerning the information they remember at that time. Detailed information about the scenario and the modifications is given in Appendix A.

As Çak (2011) mentioned, the scenario consisted of high and low workload segments and SA queries were asked during the high workload segment whose timings were specially selected and validated according to pilot studies by SMEs. Eight online queries and thirteen offline queries were administrated. Online queries were carried out orally while offline queries were asked in two sessions, with five and eight questions at a time. As a modification to Çak's study, offline queries were not formatted as multiple choice questions (Appendix A). Since each flight has characteristics of its own, the predefined answers are generally inadequate to capture specific conditions in each flight. Correct answers with respect to the current situation of that flight were recorded by the experimenter during each flight and accuracies of the responses were calculated accordingly.

3.5.2 AOSPAN

For an assessment of working memory capacity, the Automated Operation Span (AOSPAN) test was used (Unsworth et al., 2005). This test is the computerized version of operation span task (Turner & Engle, 1989) and it is the one of most common way of doing WMC measurement in the literature. As Çak (2011) denotes, this task taps on complex working memory capacity and it is valid among different cultures and languages. With this respect, it is suitable for working memory capacity assessment with pilots. Another reason for selecting this task is challenging nature of the test. No ceiling effect was observed despite highly qualified participants. In the task, participants are expected to recall a sequence of letters, each of which are presented after a simple mathematical query. For example, the answer for the operation “(5*6)-5” is asked and after it is answered a letter is shown. This sequence happens repeatedly and at the end, all presented letters have to be recalled in order. Screenshots for this task can be seen in Figure 3.5.

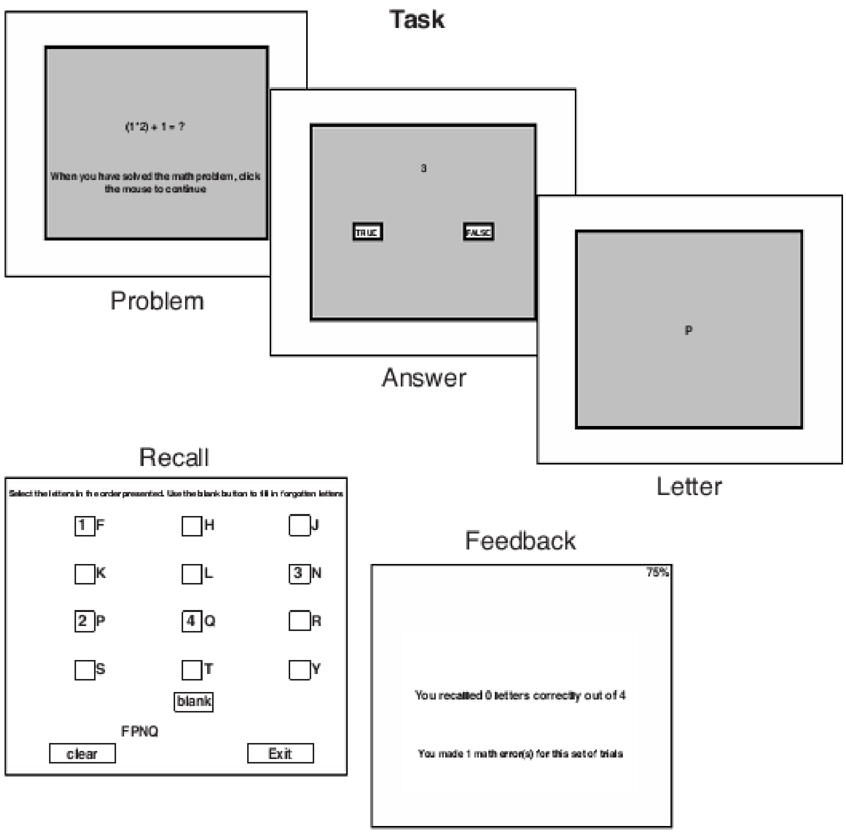


Figure 3.5 AOSPAN task screenshots adopted from Unsworth et al. (2005, p.500)

If the mathematical operations, which are displayed before the letters, are not answered correctly 85% of the time, the letter recall success is not validated. In a sequence, from three to eight letters are shown to be remembered. After the eight-letter sequence is answered, the score is calculated as the sum of perfectly recalled sequences through the experiment.

3.5.3 Stroop

To measure inhibition capability in attention, the Stroop task, an indicator of well-managed attention (MacLeod, 1991), is used. The experiment is based on the ability of preventing a habitual response in favor of the goals of the task. Participants are presented strings in five different colors; red, blue, green, yellow, and black. They are expected to name the color of the string each time. In neutral part, the string "@@@@@" is given in different colors. In the second part, congruent cases are presented where the string is the name of its color. In the last case, congruent and incongruent cases were presented randomly. 25% of the last cases were incongruent cases. For incongruent cases, the string was again a color name but different than its color. Inhibition comes into play in this case where participants have to suppress the prepotent response which is word reading.

The difference in response times between congruent and incongruent cases was used as the inhibition delay[18] as commonly used in the literature (MacLeod, 1991). In response time calculation, any wrong color namings during the incongruent cases has been excluded.

3.5.4 Coşkunöz Visual Attention Test

Divided attention is measured by a dual visual task, Coşkunöz Visual Attention Test developed by Er, Sümer, Koku, Mısırlısoy, Coşkan, Erol-Korkmaz, Sümer, Ayvaşık and Eriş (2011) in which participants are expected to follow and respond to two visual tasks running at the left and the right sides of the screen at the same time. On the left side, a red dot travels through the borders of a hidden shape without leaving a trace. When it is finished, participants are asked to find this shape among five alternatives. On the right side, 4 drawings of a tool were presented, one of which is slightly different. The different one was expected to be selected. The task on the left

[18] For inhibition delay calculation, Çak used RT difference between incongruent and neutral cases. Another inhibition delay calculation was done with this method and it was used in the comparison with Çak's study

side can be considered as the primary task, since, as it runs, the task on the right side (the secondary task) runs for 5 to 7 times. The divided attention capability was measured by the combined score which is the number of the correct answers for the secondary task that were answered in the period where the primary task was correctly answered. Figure 3.6 figure 3.7 shows the experiment screenshots.

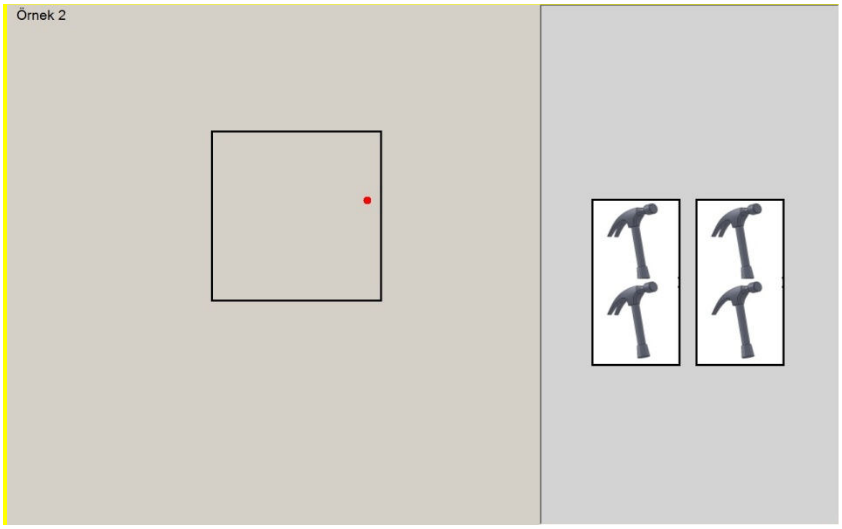


Figure 3.6: Coşkunöz Visual Attention Test (left: red dot travelling & right: awaiting answer)

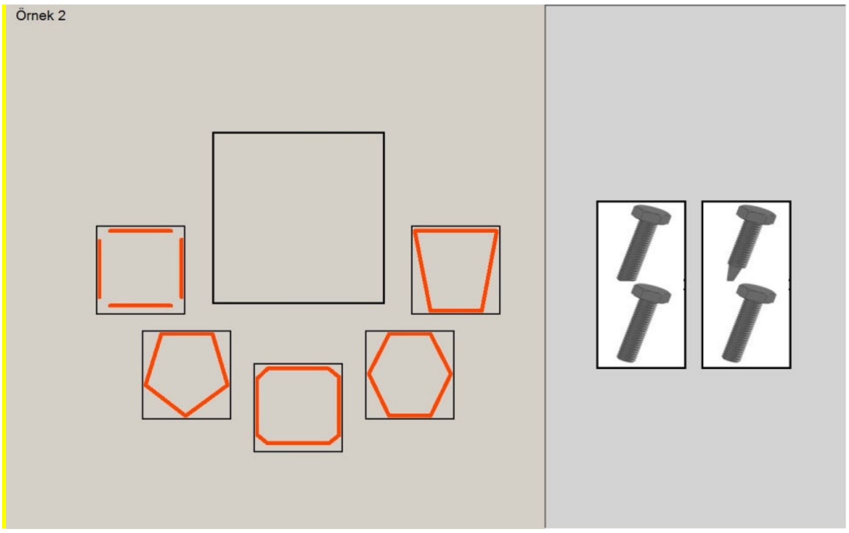


Figure 3.7: Coşkunöz Visual Attention Test (awaiting answer on both sides)

3.6 Procedure

For the simulated flight task, all participants with different expertise levels are expected to complete a pre-flight at home. Since the scenario in the experiment includes novel events, several equipment failures and special navigation points, it was aimed that participants would go over their skills for these special cases and would have a chance of practice. Detailed information about the pre-flight scenario can be found in Appendix A.3.

On the experiment day, all participants firstly read and sign the Informed Consent Form which is given in Appendix D. After that, they read the test instructions (Appendix E) and they are further informed about the general course of the flight by the experimenter. In Çak's study video recording was done during the flight. However in this study, only voice recording has been done for the purpose of acquiring response time for online queries. Not any required information was lost due to this change.

During the flight, for the assessment of SA two types of queries have been used. For online queries, flights were not paused and participants answered the questions as fast as possible while they keep piloting. For offline queries, on the other hand, the experimenter paused the flights and shut down the monitor of the simulation computer. In this case, participants did not have any time constraint but they had to answer the questions only with the information they remember at that time without looking at the simulation computer or the flight charts.

The simulated flight task took 75 minutes in average. After the flight, participants were asked to fill the workload survey. In this survey, information about the workload of the flight phases, difficulty and complexity of the flight, personal motivation and performance were collected by self-assessments of the participants. The complete workload survey is given in Appendix F.

The rest of the behavioral measurements consisted of the cognitive capacity tests; AOSPAN, Stroop and Coşkunöz Visual Attention Test which were executed in a counterbalanced sequence. After a fifteen minute break, participants took these tasks on the simulation computer and required instructions were provided by the task software. During the Stroop task, responses of participants are checked by the experimenter on a list and wrong answers are excluded from response time calculations. In average, AOSPAN task took 25 minutes, Stroop task took 7 minutes and Coşkunöz Visual Attention Test task took 5 minutes.

The behavioral measurements were finished after the experimenter handed participants the debriefing sheet (Appendix G).

fNIR brain imaging study has been conducted at METU Informatics Institute with the participants from Ankara. The optical brain imaging setup could not be transferred to other cities. However, recordings from 10 participants have been completed and the amount of participants is considered to be enough for this type of study (Friston, Holmes and Worsley, 1999). Brain activation measures of participants have been collected during the simulated flight task. In the beginning of each session, a relaxation period of about 1 minute has been recorded as a baseline for the oxygenation level in the brain. For this period participants were asked to rest with their eyes closed. Recordings from this period are later used for determination of the relative brain activity during the tasks. SPSS version 20 is used for statistical analysis of the data.

3.7 The Current Study

This thesis investigates the cognitive basis of the phenomenon of SA in terms of individual differences in working memory capacity, inhibition delay and divided attention capacity together with the effects of expertise. At the end of behavioral experiments, variances in SA scores and SA RT measurements are expected to be predicted by the variances by the factors above. Therefore, it has been hypothesized that:

H1. Working memory capacity is a positive predictor of SA scores.

H2.Working memory capacity is a negative predictor of SA RT measurements.

H3. Inhibition delay is a negative predictor of SA scores.

H4. Inhibition delay is a positive predictor of SA RT measurements.

H5. Divided attention capacity is a positive predictor of SA scores.

H6. Divided attention capacity is a negative predictor of SA RT measurements.

H7. Expertise is a positive predictor of SA scores.

H8. Expertise is a negative predictor of SA RT measurements.

As the second part of the study, mental workload is investigated in terms of participant declarations and fNIR measurements. Effects of the difficulty in the simulated flight task are expected to be observed in both subjective reports and brain activation measurements by fNIR technique.

H9. As the difficulty of flight increases, there is an increase in perceived workload.

H10. As the difficulty of flight increases, there is an increase in oxygenation at the prefrontal cortex.

CHAPTER 4

RESULTS

In this chapter, results of the study will be given. Obtained data can be found in Appendix H. Results from the simulated flight task are given as “Online”, “OnlineRT”, “Offline”, and “Combined“, which stand for Online SA scores, Online SA RT measurements, Offline SA scores, and Combined SA scores, respectively. WMC represents AOSPAN task score, STROOP is the inhibition delay obtained from Stroop task and DIVATT is the combined score for Coşkunöz visual attention test.

4.1 Behavioral Tasks

At the end of behavioral tasks several data have been obtained. From the Simulated Flight task, Offline SA scores, Online SA scores, Online Response Times (RT) and Combined scores have been recorded. Offline SA score is the sum of scores from offline SA queries. Online SA score is the sum of scores from online SA queries. Online SA RT values have been obtained by summing up RTs for successfully answered online queries. Combined Score, on the other hand, is just the summation of Offline and Online SA scores. Details of SA queries and scoring can be found in Appendix A. From the AOSPAN task, Operational Span (OSPAN) score is used (Unsworth et al., 2005). This score is the sum of number of letters that are perfectly recalled in the given sequence. Performance in Stroop task is calculated by the average delay in RT between incongruent and congruent cases. This delay is regarded as inhibition delay. Lastly, the performance in Coşkunöz visual attention task is calculated as the number of correct answers in the secondary task (selecting the different tool among four on right side of the screen) given in the period of successfully answered primary tasks (selecting the shape that the red dot travels on the left side of the screen). Details of scoring for all tasks can be found in Chapter 3.

For expertise, participants' total flight hours on several simulation platforms have been used.

For the purpose of finding contributions of predictors to SA findings, expertise (EXP), working memory capacity (WMC), Stroop (STROOP) and Coşkunöz Visual Attention Test (DIVATT) scores have been used in linear multiple regression analyses. Four regression analyses for Offline SA, Online SA, Online RT and Combined SA have been done on SPSS Version 20. Data from thirty-five participants are used.

Obtained data has been analyzed for descriptive statistics in the beginning. OnlineRT, EXP and STROOP values have non-normal distributions as given in Appendix H. Due to the non-linear relations observed between EXP and dependent variables, log transformation is applied on expertise and the new EXP_L variable is obtained.

Correlation analysis has also been performed to see the connections between the variables as given in Table 4.2. Combined scores have significant correlations with Online and Offline Scores since it is calculated as a sum of the two (r= .724, $p < .00$ and r=.821, $p < .00$, respectively). Online RT, Offline and Combined Scores are significantly correlated with one of the predictors, STROOP (r= .498, $p < .01$, r= -.549, $p < .01$ and r=-.567, $p < .00$, respectively) and Offline and Combined Scores are significantly correlated with EXP_L (r= .492, $p < .005$, and r=.430, $p < .05$, respectively). Correlations between SA scores and STROOP are negative since inhibition delay which represents the delay in the incongruent cases is a measure of lack of inhibition capability. Among the predictors, no correlation has been found.

Table 4.1 *Correlations among variables*

Correlations

		Online	OnlineRT	Offline	Combined	WMC	STROOP	DIVATT	EXP_L
Online	Pearson Correlation	1	,008	,200	**,724****	,008	-,309	,025	,143
	Sig. (2-tailed)		,963	,248	,000	,965	,071	,885	,411
	N	35	35	35	35	35	35	35	35
OnlineRT	Pearson Correlation	,008	1	-,283	-,195	,239	**,498****	-,103	-,221
	Sig. (2-tailed)	,963		,099	,262	,166	,002	,555	,203
	N	35	35	35	35	35	35	35	35
Offline	Pearson Correlation	,200	-,283	1	**,821****	-,002	**-,549****	,119	**,492****
	Sig. (2-tailed)	,248	,099		,000	,990	,001	,496	,003
	N	35	35	35	35	35	35	35	35
Combined	Pearson Correlation	**,724****	-,195	**,821****	1	,003	**-,567****	,099	**,430****
	Sig. (2-tailed)	,000	,262	,000		,986	,000	,573	,010
	N	35	35	35	35	35	35	35	35
WMC	Pearson Correlation	,008	,239	-,002	,003	1	,272	,157	-,082
	Sig. (2-tailed)	,965	,166	,990	,986		,115	,368	,641
	N	35	35	35	35	35	35	35	35
STROOP	Pearson Correlation	-,309	**,498****	**-,549****	**-,567****	,272	1	-,217	-,269
	Sig. (2-tailed)	,071	,002	,001	,000	,115		,212	,118
	N	35	35	35	35	35	35	35	35
DIVATT	Pearson Correlation	,025	-,103	,119	,099	,157	-,217	1	-,018
	Sig. (2-tailed)	,885	,555	,496	,573	,368	,212		,917
	N	35	35	35	35	35	35	35	35
EXP_L	Pearson Correlation	,143	-,221	**,492****	**,430****	-,082	-,269	-,018	1
	Sig. (2-tailed)	,411	,203	,003	,010	,641	,118	,917	
	N	35	35	35	35	35	35	35	35

**. Correlation is significant at the 0.01 level (2-tailed).

Following the results from Çak (2011), hierarchical technique was used for Offline SA in order of WMC, EXP_L, other predictors and for Online RT in order of EXP, STROOP and DIVATT. It was expected to find similar predictor contributions with Çak's study. However, these orders used in the hierarchical method did not give any better results compared to Enter method. It is observed that STROOP and EXP_L were the only successful predictors. Other predictors (WMC and DIVATT) do not have any significant contribution to the variance explained and addition of them in hierarchical regression decreases the amount of variance explained by the models. Results obtained by the Enter method are given below.

For the prediction of Online SA scores, four predictor variables (WMC, STROOP, DIVATT and EXP) did not lead to any successful regression. Further details of the regression analysis can be found in Appendix H.2.

For the prediction of Online SA RT measurements, four predictor variables (WMC, STROOP, DIVATT and EXP_L) produced an adjusted R^2 of .17 ($F(4,34) = 2.756$, $p = .05$) given in Appendix H.3. This is a weak explanation on Online SA RT measurements. The only successful predictor for this regression is inhibition delay (STROOP) with $\beta = .434$, $t(30) = 2.481$, $p < .05$. Further details of the regression analysis can be found in Appendix H.3.

For the prediction of Offline SA scores, four predictor variables (WMC, STROOP, DIVATT and EXP_L) produced an adjusted R^2 of .38 ($F(4,34) = 6.233$, $p < .01$) given in Appendix H.4. This is the strongest explanation found in the scope this study. The successful predictors for this regression are inhibition delay (STROOP) with $\beta = -.495$, $t(30) = -3.270$, $p < .005$ and expertise (EXP_L) with $\beta = -.372$, $t(30) = 2.649$, $p < .05$. Since the inhibition delay is calculated as the delay in the incongruent cases, participants with better inhibition abilities had lower inhibition delays. Consequently, beta value of STROOP for the prediction of Offline SA scores is negative. Further details of the regression analysis can be found in Appendix H.4.

For the prediction of Combined SA scores, four predictor variables (WMC, STROOP, DIVATT and EXP_L) produced an adjusted R^2 of .36 ($F(4,34) = 5.732$, *p*< .005) given in Appendix H.5. Successful predictors for this regression is the inhibition delay (STROOP) with $\beta = -.545$, $t(30) = -3.536$, $p < .05$. Again, negative beta value is obtained due to the negative relation between Combined SA score and the predictor STROOP. The other successful predictor is EXP_L with $\beta = .297$, $t(30) = 2.077$, $p < .05$. Further details of the regression analysis can be found in Appendix H.5.

The hypothesis “Working memory capacity is a positive predictor of SA scores” (H1) is not supported by the regression analysis.

The hypothesis “Working memory capacity is a negative predictor of SA RT measurements” (H2) is not supported by the regression analysis.

The hypothesis “Inhibition delay is a negative predictor of SA scores” (H3) is partially supported by the regression analysis. For Offline and Combined SA scores inhibition delay is an important predictor, while it is not predictive for Online SA scores.

The hypothesis “Inhibition delay is a positive predictor of SA RT measurements” (H4) is supported after the regression analysis.

The hypothesis “Divided attention capacity is a positive predictor of SA scores (H5) is not supported by the regression analysis.

The hypothesis “Divided attention capacity is a negative predictor of SA RT measurements” (H6) is not supported by the regression analysis.

The hypothesis "Expertise is a positive predictor of SA scores" (H7) is partially supported by the regression analysis. For Offline and Combined SA scores expertise is an important predictor, while it is not predictive for Online SA scores.

The hypothesis "Expertise is a negative predictor of SA RT measurements" (H8) is not supported by the regression analysis.

4.2 Workload Survey

The scenario was supposed to be challenging in order to investigate SA under critical conditions. Below listed assessments showed that this aim was achieved. Participants' self-assessments clearly showed that specific parts of the flight had high workload where very high amounts of attention were demanded. fNIR data were also segmented in accordance with the design of the scenario and relatedly in accordance with self-assesment results from the workload survey.

At the end of the experiment, all participants were asked to fill out the workload survey that can be found in Appendix F. This survey was formed by Hart et al. (1984) and used by Çak (2011) as well. It consisted of questions aiming to assess the participants' ideas about the simulated flight, their performance and motivation, in general.

One of the important questions was workload during the flight. Participants, on average, marked 4.4 out of 6 for overall workload during the flight. This score can be interpreted as a moderate level (Hart et al., 1984). The survey included also partial assessments of the flight phases in terms of workload. As the flight conditions got worse due to weather conditions and equipment failures, participants' responses for the workload questions also got higher. On average, they marked increasingly for workload in the partial assessments. In the beginning, take-off phase was rated as 1.6 over 6, followed by the easy flight part between LTBA and BKZ transmitters where they marked on average 1.91 out of 6. After the easy part, bad weather conditions were applied and between BKZ and YAA transmitters it was progressively scored as 2.91 over 6. The next two segments where several equipment failures were introduced under poor weather conditions have been reported to have high workload demands. Up to the intersection point ERMAN, participants reported 4.31 out of 6 and for the section including the hold at ERMAN they gave 4.6 points over 6. These assessments by the participants showed that the scenario was increasingly difficult and demanding as it had been planned. For the last section, landing phase, participants marked 4.09 points out of 6. In this part of the flight weather conditions got better and flight became easier. Consequently, workload scores given by the participants decreased. Perceived workload values are given in Figure 4.1.

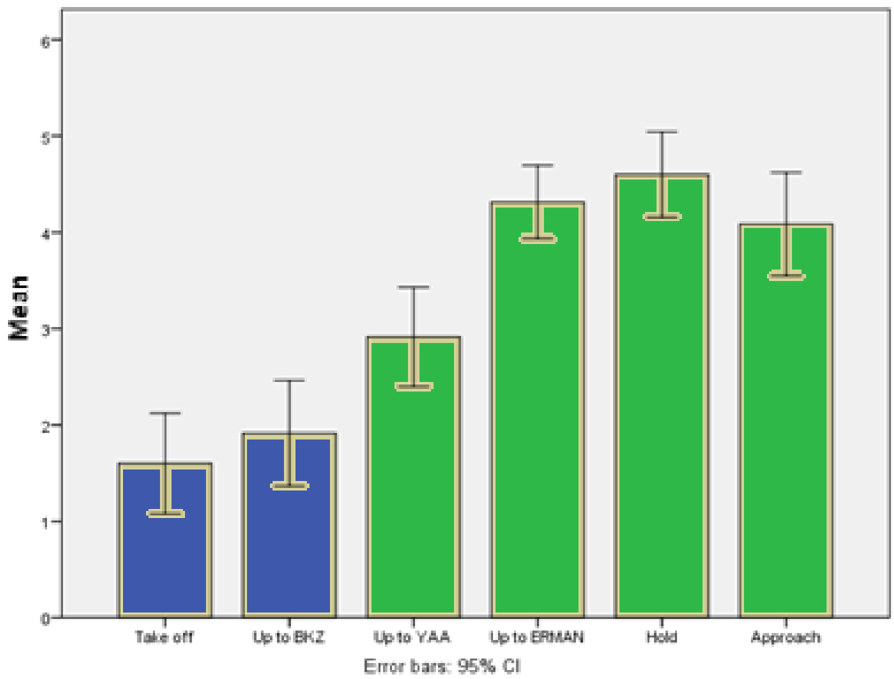


Figure 4.1 Average Perceived Workload values for phases of the simulated flight

Workload assessments were not the only questions in the survey. Participants' evaluation for stress during the flight was rated on average 3.3 out of 6 and at the end of the flight fatigue was rated on average 2.6 out of 6. It can be inferred then, that flight conditions reasonably affected the participants. Other questions listed below exhibits the challenging features of the scenario. Participants evaluated the demands of attention very high with 5,3 over 6 and the complexity of the scenario with 4,14 out of 6. They marked 4,3 for the difficulty of the flight and for their own performance 3,1 out of 6 in average. Both mental effort and activity during the flights were evaluated high as 4,4 out 6 by the participants.

As planned in the scenario, the simulated flight can be evaluated in two segments in terms of workload. Low workload part starts from the beginning till the first SA query which comes right after BKZ point. The rest of the flight constitutes high workload segment. These two segments are compared for perceived workload values with paired-samples t-tests. Results showed that mean WL declarations significantly differ between low WL (M =1.78, SD =1.32) and high WL (M =3.98, SD =1.05) segments; $t(34) = -10.70$, $p < .001$. Further details of the statistical analysis can be found in Appendix H.6. Figure 4.2 shows the average values for two segments with 95 % confidence intervals.

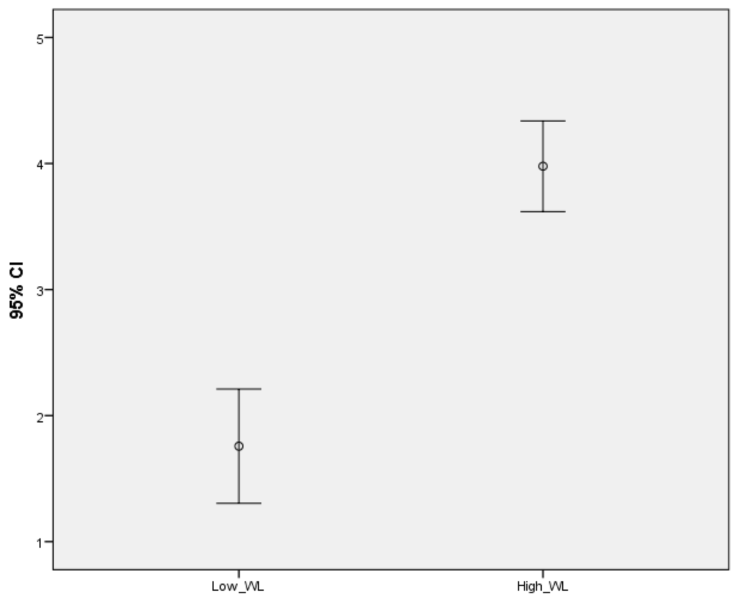


Figure 4.2: Average Perceived Workload for Low WL and High WL segments

The hypothesis “As the difficulty of flight increases, there is an increase in perceived workload” (H9) is supported after the analysis.

4.3 fNIR Study

A brain imaging study with functional near infrared (fNIR) spectroscopy has been carried out during the simulated flight task. Ten participants were tested while they were under specific task conditions. This sample size is considered to be enough for the exploratory purposes of this study (Friston, Holmes and Worsley, 1999). Brain activations in terms of oxygenation have been recorded and analyzed to investigate neural correlates of mental workload during a simulated flight.

Oxygenation values for each voxel in the sensor area were recorded during different blocks of the task. The simulated flight task has been evaluated in two blocks that are "Low WL" (Workload) and "High WL". Low and High WL distinction in scenario was planned during the formation of the scenario. In order to provide familiarization period in the beginning of the flight, a low WL part was introduced. During this part, everything was planned to happen as expected and participants started to feel comfortable about the environment. With the administration of the first online query, flight conditions were set to get worse with several equipment failures. Difficult navigation problems were also introduced as explained in Chapter 3. Consequently, two different WL segments have been formed during the flight and this condition was also approved by the Workload survey answered by the participants.

For an overall analysis of the data, oxygenation values from each participant have been averaged and data representing fourteen voxels[19] were compared using Low and High WL segments as explained above. Related bar-chart is given in the Figure 4.2 below. It is clearly observed that for most of the voxels there is a considerable difference between Low WL and High LW conditions. This finding also supported the theoretical Low and High WL distinction in the scenario. Participants had more brain activation during High WL segment as it was expected. Results from the workload survey also corroborated with the fNIR findings.

[19] 2 channels that correspond to voxels 11 and 13 were found to be defective, and hence removed from further analysis. Analysis of the fNIR data was conducted on the remaining 14 channels.

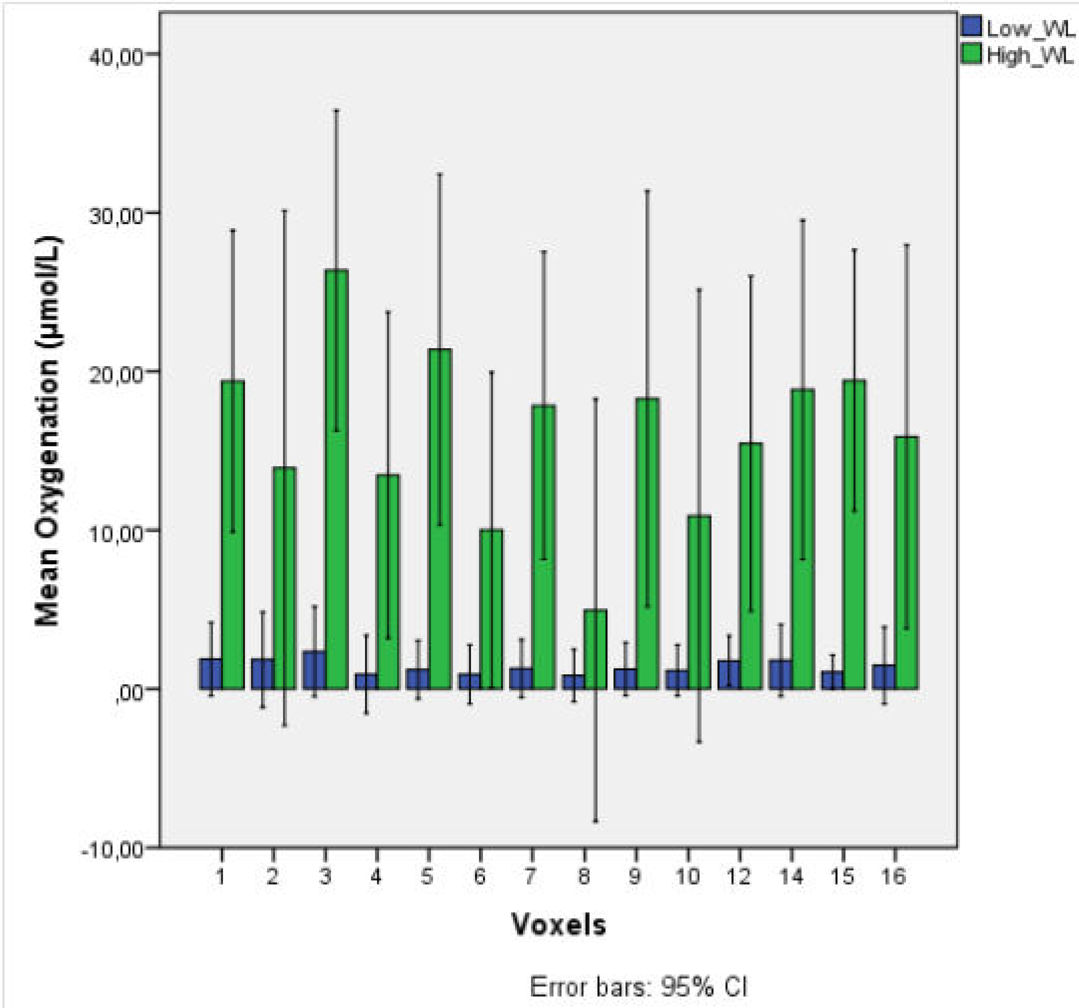


Figure 4.3: Average Oxygenation Data from all participants for Low WL and High WL segments

fNIR signals from each voxel were compared in Low WL and High WL conditions with paired-samples t-tests. Significant difference has been observed in all voxels except for voxels 2,8, and 10. Related significance values for voxels are given as: voxel 1 (t(9) =-5.096, p < .005), voxel 3 (t(9) =-6.589, p < .001), voxel 4 (t(8) =-3.626, p < .01), voxel 5 (t(8) =-4.804, p < .005), voxel 6 (t(8) =-2.425, p < .05), voxel 7 (t(8) =-4.723, p < .005), voxel 9 (t(8) =-3.332, p < .05), voxel 12 (t(9) =-3.230, p < .05), voxel 14 (t(8) =-4.466, p < .005), voxel 15 (t(8) =-5.481, p < .005), and voxel 16 (t(8) =-3.379, p < .05). Results showed that percieved workload reported by subjects was clearly observed in fNIR measurements. Further details of the statistical analyses can be found in Appendix H.7.

The hypothesis "As the difficulty of flight increases, there is an increase in oxygenation at the prefrontal cortex." (H10) is supported after the analysis.

4.4 Comparison with Çak's Study

As mentioned before, this study has been constructed over Serkan Çak's doctoral studies (2011). Çak has conducted simulated flights on a full flight simulator and cognitive capacity tests, which are replicated in this study. However, in the present study participants were not professional pilots but simulator pilots. The same simulated flight task has been used with minor modifications as given in Appendix A. Cognitive capacity tasks, AOSPAN and Stroop have been carried out in with no modifications. Lastly, a different task for assessment of divided attention has been used in this study due to differences in professional and simulator piloting practices. Çak has also carried out eye-tracking experiments with professional pilots which were not replicated in this study.

Çak's study aimed to investigate the relationship between SA with individual cognitive differences. SA measures were taken and tried to be explained by working memory and attention measures. At the end of his studies, 58% of variability in offline SA measures was accounted for by variances in working memory and expertise measures. In his analysis, WMC was the most successful predictor (β = .675, $t(31)$= 5.31, p<.00), whereas the other predictor, expertise (β = .278, $t(31)$= 2.35, p<.05) was not that successful. For online SA measures (average RTs for correct answers in online queries), 52% of variability was accounted for by variances in inhibition, divided attention and expertise measures. The predictors in order of strength are listed as expertise (β= -.470, $t(31)$= -3.73, p<.001), divided attention (β= .313, $t(31)$= 2.25, p<.05) and inhibition (β = .260, $t(31)$= 2.058, p<.05). One important thing to note here is that in Çak's study pilots' expertise has been evaluated by not flight hours of the real flight but flight hours of the full flight simulator that the experiment was administered. For simulator pilots, on the other hand, their simulated flight experiences on different platforms were evaluated totally.

A comparison between the results of this study and Çak's study has been performed using the common questions and tests in both of the experiments. Offline SA scores and online SA RT averages have been recalculated in the way Çak did due to the different accuracy assessments applied in the two studies. Only five online queries which were same in both studies were taken into account. For the comparison, inhibition delay was recalculated as Çak did. Consequently, Offline, OnlineRT, EXP, WMC and STROOP variables were analyzed using independent group t-tests. Results showed that there are no significant differences between groups except for the Offline SA scores. Offline SA scores from professional pilots (M=720.37, SD=195.8) and simulator pilots (M=602.86, SD=166.6) were significantly different from each other; $t(68)=2.704$, $p<.01$. Further details of the statistical analysis can be found in Appendix H.8.

This results show that despite the similarities in cognitive capacity tests and expertise, professional pilots were distinctively more successful in offline SA queries. The differences in pilot training backgrounds and practice systems structures are candidate reasons to explain this finding. A possible difference in online SA scores between professional and simulator pilots could not have been observed since in Çak's study online SA scores are not evaluated.

CHAPTER 5

DISCUSSION

5.1 Behavioral Tasks

This study aimed to find meaningful relations between the concept of SA and individual cognitive differences. For assessment of SA, two types of queries have been administered during the simulated flight task; online and offline queries. Online queries were answered as the simulation runs and accuracies of the answers and RTs were recorded. Offline queries, on the other hand, were asked after the simulation was frozen, and flight screen and flight charts were blacked out. Accuracies of answers were recorded. Consequently, Online SA scores, Online RT measurements, Offline SA scores and Combined SA scores were obtained. Cognitive capacity tests were also carried out after the simulated flight to measure WMC, inhibition and divided attention capabilities. WMC scores have been acquired as operational span scores from AOSPAN task. STROOP scores were calculated as the average RT delay between incongruent and congruent cases in Stroop task. DIVATT (Divided Attention) scores were the results of Coşkunöz visual attention test. Lastly, expertise (EXP_L) for simulator pilots has been determined using logarithmic transformation of total simulation flight hours.

Considering the whole regression results, first unexpected finding to examine is the absence of expertise as a predictor for online SA scores and RT measures. The regression results report that expertise contribute to the explanation of the variances in only offline SA and combined SA scores. This is against the basic notion that experienced pilots would be more competent in all aspects of flight considering SA performance. Current results could be interpreted in two different ways. The first comment is that SA performance is not strictly linked to flight performance and relatedly not strictly linked to expertise, as well. However, as described in the Introduction chapter, SA is regarded as a critical component of the flight performance. It is reported that SA is operationally very important and inadequate SA causes the 88% of all human errors (Endsley, 1995b). Here, a conceptual problem is revealed. The concept of SA that is given as the reason of human errors and the other concept of SA that is obtained from query answers are probably not the same thing. When experts inspect an accident, they notice a failure as the reason of the accident that the pilots (or operators) miss an important piece of information. This is regarded as the “loss of SA”. However, in the assessment of SA (by different measurement techniques) pilots or operators are questioned for some information, which is not necessarily crucial at the moment of the questioning. Consequently, “the actual SA”, loss of which found as a strong reason for accidents is not necessarily same as “the measured SA”. These two concepts definitely share some portion of operator capabilities but it is not strongly defendable to claim that they are the same concept. Vaitkunas-Kalita, Landry, and Yoo (2011) also pointed a similar inconsistency in the folk and scientific uses of the term SA. They have investigated 81378 reports from the Aviation Safety Reporting System where pilots and controllers report incidents. The usage of the term SA in reports is compared to its usage in scientific research in terms its reference and content. Only 1.4 % of reports were found to have the same meaning with the term SA employed by researchers in this area. References to the term SA from the aviation professionals and researchers do not necessarily point to the same phenomenon. To sum up, there are differences in the references to the term SA and expertise may be linked to the former concept of “actual SA”, but less likely to be connected to the latter concept of “measured SA” which is obtained in the scope of this thesis.

Second comment to interpret the inadequacy of expertise as a consistent predictor in this study is the difficulty of assessment of expertise in simulator piloting. In Çak’s study, professional pilots’ flight hours on the test simulator are taken as the measure of their levels of expertise. Similarly, in this study the simulator pilots’ expertise levels are determined in terms of flight hours, but those figures are based on the pilots’ personal declarations. Despite the fact that most of the participants are simulator pilots from virtual airlines, which have standards for training and expertise levels, unfortunately personal declarations may contain simulated flights from different platforms with different setups. Another reason for the difficulty of assessment of expertise in simulator piloting is that the simulated flights that the pilots execute do not have specific difficulty levels. Again in the virtual airlines, training procedures are followed for the recruitment of simulator pilots, but simulator pilots do not take further training as their ranks (relatedly experience) increase. They might prefer to fly under easy conditions as their personal choice or want to improve their skills with increasingly difficult flights. After all, there are reasons to suspect the validity of flight hours as a measure that is representative of expertise in this study.

Another important issue for this study is that WMC is found to be not explaining any of the variances in the SA measurements. Working memory is considered to have a central importance for SA (Durso & Gronlund, 1999; Endsley, 1995b). Tasks, systems and timely information that are critical for SA are kept and processed by working memory (Wicken, 1999). Also, there are several studies that the correlation between SA measures and WMC is given (Durso et al., 2006; Gonzales and Wimisberg, 2007). However, in this study, no correlations or no similarities in variances has been observed between WMC and SA measures. A possible explanation to this finding is the vast range of differences in simulator piloting practices compared to professional piloting. Simulator pilots are generally self-educated and unexpectedly have their own way of piloting due to the lack of formal education. During the experiments, it is observed that they were likely to use autopilots and automatic navigation devices despite they were not allowed during the simulated flight task. It is probably because they do not go through formal education steps and develop their own practices. In professional piloting, on the other hand, every pilot starts to fly with basic aircrafts which are not equipped with autopilots or navigation devices. This needs more control effort and more attention on the elements of flight due to lack of these assisting equipments. In this way, professional pilots develop an inner understanding of the flight which is less likely to develop in simulator pilots. Consequently, it is revealed that simulator pilots' performances on SA measurements are not determined by the systematical factor of working memory capacity, but possibly determined by individual self-training differences.

Following the discussion on lack of a formal education and an inner understanding of the flight, the results for the predictors STROOP and DIVATT seem to lose their importance. However, even under these considerations, it is important to note that STROOP happened to be the consistent predictor for Online RT, Offline and Combined SA measurements. Attention control capability captured by Stroop task, unlike utilization of working memory for elements of flight, is found to be effective in SA performance. Since the SA measurements were carried out as the participants were busy with piloting, answering these queries required handling both of the tasks. At this point, attention control became an important capability. Possibly with this connection, STROOP turned out to be a consistent predictor. Along similar lines, the reason why DIVATT did not turn out to be a good predictor can be explained. Compared to the real flight situation, simulation environment is simple in terms of environmental factors. Simulation environment consisted of a PC and a joystick while real flight contains two environments, inside and outside the airplane. Professional pilots observe both the equipments inside the plane and weather conditions outside the plane. Consequently, it might be proposed that divided attention capacity for simulator pilots is not as important as it is for professional pilots. Nevertheless, due to the effects of unsystematic differences in simulator pilots' practices as mentioned above, these comments have to be considered cautiously.

5.2 fNIRS Study

Optical brain imaging measurements with fNIR during the simulated flight task showed that as the perceived workload increased, oxygenation in the prefrontal cortex was also increased. This result is compatible with several other studies assessing the effects of workload with PET and fMRI techniques (Badre and Wagner, 2007; Cohen et al., 1997; Osaka et al., 2007). Also other fNIR studies report similar results (Ayaz et al, 2010; Izzetoglu et al., 2004; Schreppel et al., 2008). Despite the less complicated structure compared to other brain imaging techniques, fNIR technique has successfully differentiated the workload segments. fNIR studies in literature generally evaluate one or two voxel readings that have significant results. However, current study does not focus on any specific voxel. Oxygenation values from all prefrontal cortex is evaluated and significant increases in oxygenation connected to task difficulty were observed in 11 out of 14 voxel readings in the overall assessment. These results are also in line with perceived workload declarations. Usage of fNIR brain imaging technique in the assessment of mental workload is found to be appropriate in connection with the above listed studies. Moreover, the difference observed between different flight episodes were more pronounced in some voxels as compared to others. A follow up study with more subjects may investigate how such differences are associated with activation patterns observed during tasks such as AOSPAN, Stroop and DIVATT to better account for the cognitive constructs underlying SA.

5.3 Comparison with Çak's Study

In the beginning of this study, it was planned to have comparable research with Çak (2011) such that similarities and differences between simulator and professional pilots would be analyzed. SA measurements were expected to be parallel in two groups to investigate further effects of individual cognitive differences. However, significant differences were observed in Offline SA measurements as given in Results part. There were no significant differences in cognitive capacity measures. Simulator pilots having similar cognitive capacity measures with professional pilots were clearly less successful in SA measurements.

This result shows that differences between professional and simulator pilots are more foundational than we thought. Simulator pilots generally train themselves using several online resources and it is far from being standard and systematic. Compared to the training in professional piloting, it is rather simple. Practices that simulator pilots follow are in accordance with professional pilots; nevertheless they are not supervised and corrected rigidly during the training period. As noted above, it was observed that each simulator pilot develops his own way of controlling the airplane. They differ in ways of using flight equipments, steering the airplane and following procedures. This vast range of uncommon practices shows the results of self-training. Unlike the training in professional piloting, simulator pilots start their training with big jet airplanes with automatic controls for navigation and even for landing. They do not have to control altitude, airspeed and heading of the airplane at the same time. Due to the practice of using automated controls, it is likely that, simulator pilots fail to develop basic skills in professional piloting. Professional pilots, on the other hand, fly with basic airplanes during their training where they have to deal with all of the above listed issues. Even if two groups were tested on the same scenario with common events, foundational differences cause different results in SA measurements. Unfortunately, there is currently no literature on these differences.

Another reason of differences with Çak's study is the assessment of expertise. In Çak's studies, professional pilots' expertise was determined as the number of flight hours spent in the full flight simulator, not of the real flight. Since all participants in his study were professional pilots with systematic training backgrounds, flights on the full flight simulator enable them to overcome familiarization problems. This was a clear advantage for Çak's study that most of the participants had previous experiences and related flight hours for the full flight simulator on which the experiment was done. Only seven out of thirty six participants had less than 20 hours of experience on the full flight simulator. However, assessment of expertise was problematic for simulator pilots. Experiences from flights on different flight platforms with several setups were expected to be represented by total flight hours. Consequently, total flight hours were used as expertise level. Most of the participants were pilots from virtual airlines which have some standards but not at a professional level. Simulator pilots are accepted in these communities after completing some specific flights. After joining the virtual airline, simulator pilots can fly on different levels of difficulty at their own disposal. Due to lack of standards and strict procedures in simulator piloting, total flight hours happened to be a poor indicator of expertise. As a result, this assessment technique for expertise was not successful for prediction of SA measurements.

Considering the notions above, it can be concluded that there are fundamental differences in professional and simulator piloting. Training backgrounds, operation environments and experience development are prominent reasons for these differences. However, to our best knowledge, there are no studies in the literature specifically focusing on the comparison of these two groups. It has been observed that, for these two groups with different training and practice backgrounds, similar cognitive individual differences do not result in similar SA scores for the same flight task. For simulator pilots, it is revealed that there are different patterns of relationships between SA and individual cognitive differences.

5.4 Limitations

For this study, finding simulator pilots to participate was a challenge. At the end, behavioral tasks were completed with 35 participants and statistical analysis was carried out. Results of the regression analysis should be evaluated carefully due to limited number of participants.

As noted in Çak's study (2011), Online RT values were not easy to exactly calculate due to hesitations and murmuring in the answers. During Online SA queries, last answers given by the participants were taken into consideration with related RT measurements.

This study has been carried out in four different cities. Experiment conditions were tried to be equated by using the same setup in each location and doing the experiments in isolated office environments. Nevertheless, unnoticed details about the environments might have affected the results although no such conditions exist to the best of our knowledge.

Assessment of individual cognitive differences was done by capacity tests. These tests, namely, AOSPAN, Stroop and Coşkunöz visual attention test, come with their own limitations in representation of cognitive concepts. It is important to note that obtained results are subject to change if the experiments were replicated with different cognitive capacity tests.

Expertise assessment for simulator pilots was an important drawback as noted before. Since all the simulator pilots do not use the same simulation environment, experiences have been determined by the participants' declarations of their total simulation flight hours from different platforms. However, these platforms have varying features and it is very difficult to find an evaluation criterion for expertise from different platforms. Another limitation is the lack of a standard training background for simulator pilots. Simulator pilots who are attending online airline flights have to complete some training steps but these are not systematic as in professional piloting. To overcome this problem a pre-flight was expected to be completed before the experiments. However, still inadequate proficiency in simulator piloting of some participants might have affected the results.

5.6 Future Work

Visiting back the discussion of "the actual SA" and "the measured SA", assessment of such a complex cognitive construct is very challenging. It has been revealed that there are still problems in defining and operationalizing SA. Research communities have offered several conceptualizations (Breton & Rousseau, 2001) while those conceptualizations do not reflect the use of the term SA in professional aviation community (Vaitkunas-Kalita, Landry, Yoo, 2011). To overcome this problem, an improvement might have been introduced by some changes in the experiment scenarios for the assessment of SA with respect to its relationship with task performance. In order to do that, some specific task measures which are critical in the case of loss of SA could also be taken as well as SA measures. Consequently, validity of the SA measures with respect to critical task measures can be obtained. These task measures should be carefully chosen to observe the results of good or bad SA.

At the end of this study, it has been clearly seen that there are unexpectedly big differences between simulator and professional pilots. These differences are probably caused by the distinctness of training systems and environmental differences during flights. In order to be able to compare these two groups, further research on training systems and flight environments have to be carried out. Training procedures for simulator pilots should be qualitatively investigated in comparison with training procedures for professional pilots.

CHAPTER 6

CONCLUSION

In this study, behavioral experiments to investigate the relationship of SA with individual cognitive differences and brain imaging study to observe the perceived workload with fNIR data have been completed. This study extends Çak's research by exploring SA with simulator pilots. The experiments are completed in the real task environment with a challenging flight scenario. To our best knowledge, the cognitive exploration of the SA has been done with simulator pilots for the first time in the literature. It is important to note that in the scope of this study experimental assessment of SA was carried out with simulator pilots with different levels of experience under environmentally valid conditions. Behavioral experiments were composed of a simulated flight task, AOSPAN, Stroop and Coşkunöz visual attention tasks for assessment of SA, working memory, inhibition, and visual divided attention capacity, respectively. Brain imaging study, on the other hand, was carried out during the simulated flight task to explore changes in relative levels of oxygenation in the prefrontal context with respect to perceived workload.

At the end of the study, the complex cognitive process of SA was aimed to be explained by the given predictors experimentally. Hypotheses about prediction of SA were tested based on four different measures of SA, namely Online SA, Offline SA and Combined SA scores and Online SA RT. None of the predictors (expertise, working memory capacity, inhibition delay, or divided attention capacity) was found to be explaining the variance in Online SA scores. Expertise, calculated by total simulation flight hours of simulator pilots, was found to be a predictor for only Offline and Combined SA scores. Inhibition delay, measured by Stroop task, was an important predictor for all SA measurements, except for Online SA, with the biggest explanation power. Other predictors, working memory and divided attention capacity were not successful in explaining any of variances in SA measurements.

In comparison to the preceding study carried with professional pilots (Çak, 2011), results from the current study did not reflect a clear picture of a cognitive grounding for SA especially for individual cognitive differences in working memory and divided attention capacities. However, meaningful relations of SA with inhibition delay and expertise have been observed. It was discussed that SA performances of simulator pilots were affected by random individual practices rather than the systematic effects of individual capacities.

The effect of task difficulty in the simulated flight task, on the other hand, was observed in both perceived workload declarations and oxygenation measurements from the prefrontal cortex. These results showed that the planned workload distinction in the scenario was effective and this effect was also successfully captured by fNIR optical brain imaging technique.

REFERENCES

Ackerman, P. L. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin, 102*(1), 3-27.

Adams, M. J., Tenney, Y. J., & Pew, R. W. (March 1995). Situation awareness and the cognitive management of complex systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37*, 85-104(20).

Atkinson, R. C. & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K.W. Spence, & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol.2, pp. 89–195). New York: Academic Press.

Ayaz, H., Cakir, M. P., Izzetoglu, K., Curtin, A., Shewokis, P. A., Bunce, S. C., & Onaral, B. (2012). *Monitoring expertise development during simulated UAV piloting tasks using optical brain imaging.* Paper presented at the Aerospace Conference*,* Big Sky, MT.

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.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417-423.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-89). New York: Academic Press.

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia, 45*(13), 2883-2901.

Bethke, B., Valenti, M., & How, J. (2007). Cooperative vision based estimation and tracking using multiple UAVs. *Lecture Notes in Control and Information Sciences, 369*, 179-189.

Breton, R., & Rousseau, R. (2001). *Situation awareness: A review of the concept and its measurement* (Technical Report No. 2001-220). Valcartier, Canada: Defense Research and Development.

Broadbent, D. E. (1958). *Perception and communication.* Londres, Inglaterra : Pergamon.

Bunce, S. C., Izzetoglu, K., Ayaz, H., Shewokis, P., Izzetoglu, M., Pourrezaei, K., & Onaral, B. (2011). Implementation of fNIRS for monitoring levels of expertise and mental workload. *Directing the Future of Adaptive Systems: Lecture Notes in Computer Science, 6780,* 13-22.

Carretta, T. S., Perry Jr., D. C., & Ree, M. J. (1996). Prediction of situational awareness in F-15 pilots. *The International Journal of Aviation Psychology, 6*(1), 21-41.

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.

Chance, B., Zhuang, Z., UnAh, C., Alter, C., & Lipton, L. (1993). Cognition-activated low-frequency modulation of light absorption in human brain. *Proceedings of the National Academy of Sciences, 90*(8), 3770-3774.

Cherry, C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America, 25*(5), 975-979.

Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., & Smith, E. E. (1997). Temporal dynamics of brain activation during a working memory task. *Naturei 386*(6625), 604-608.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769-786.

Cooke, N. J. (2006). Human factors of remotely operated vehicles. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50*(1), 166-169.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*(01), 87-114.

Çak, S. (2011). *Effects of working memory, atention, and expertise on pilots' situation awareness.* Unpublished doctoral dissertation, the Middle East Technical University, Ankara.

Çakır, M., Şenyiğit, A., Akay, D., Ayaz, H., & İşler, V. (2012). Evaluation of UAS camera operator interfaces in a simulated task environment: An optical brain imaging approach. *Advances in Brain Inspired Cognitive Systems: Lecture Notes in Computer Science*, *7366*, 62-71.

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

Deutsch, J. & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, *70,* 80-90.

Duffy, C. M. (2001). Situation awareness analysis and measurement. *Human Factors and Ergonomics in Manufacturing & Service Industries, 11*(4), 383-384.

Durso, F. T., Bleckley, M. K., & Dattel, A. R. (2006). Does situation awareness add to the validity of cognitive tests? *Human Factors: The Journal of the Human Factors and Ergonomics Society, 48*(4), 721-733.

Durso, F. T., & Gronlund, S. D. (1999). Situation awareness. In F. T. Durso, R. Nickerson, R. Schvaneveldt, S. Dumais, S. Lindsay, & M. Chi (Ed.), *The handbook of applied cognition* (pp.283–314). Wiley.

Durso, F. T., Hackworth, C. A., Truitt, T. R., Crutchfield, J., & Nikolic, D. (1998). Situation awareness as a predictor of performance in en route air traffic controllers. *Air Traffic Control Quarterly, 6*(1), 1-20.

Durso, F. T., & Sethumadhavan, A. (2008). Situation awareness: Understanding dynamic environments. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 50* (7), 442-448.

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. *Human Factors and Ergonomics Society Annual Meeting Proceedings,32*, 97-101(5).

Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37* (20), 65-84.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37*(33), 32-64.

Endsley, M. (1997). The role of situation awareness in naturalistic decision making. In C. Zsambok and G. Klein, (Eds.), *Naturalistic Decision Making*. Lawrence Erlbaum Associates.

Endsley, M. R. (2004). Situation awareness: progress and directions. In S. Branbury & S. Tremblay (Eds.), *A cognitive approach to situation awareness: Theory and application* (pp. 317–341). Ashgate.

Endsley, M. (2006). Expertise and situation awareness. In K. A. Ericsson, N. Charness, P. Feltovich, & R. Homan, (Eds.), *The Cambridge Handbook of Expertise and Expert Performance*, (pp.636–651). New York: Cambridge University Press.

Endsley, M. R., & Bolstad, C. A. (1994). Individual differences in pilot situation awareness. *International Journal of Aviation Psychology, 4*(3), 241-264.

Engle, R. W., & Kane, M. J. (2003). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In Brian H. Ross (Ed.), *Psychology of learning and motivation* (pp. 145-199). Academic Press.

Er, N., Sümer, H. C., Koku, B., Mısırlısoy, M., Coşkan, C., Erol-Korkmaz, H. T., Sümer, N., Ayvaşık, H. B., & Eriş, A. (2011). Çoşkunöz visual attention test. Unpublished instrument.

Ericsson, K. A., & Charness, N. (1994). Expert performance, its structure and acquisition. *American Psychologist, 49*(8), 725-747.

Ericsson, K. A., & Delaney, P. F. (1999). Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance. In A. Miyake, & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 257–297). New York: Cambridge University Press.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*(2), 211-245.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology, 47*(1), 273-305.

Eriksen, W. C. & James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception and Psychophysics, 40*, 225-240.

Ericksen, W. C. & Yeh,Y. (1985). *Allocation of attention in the visual field*. Washington, DC: American Psychological Association.

Fracker, M. L. (2003). *Measures of situation awareness: Review and future directions*. Retrieved 06, 05, 2012, from U.S. Department of Commerce, National Technical Information Service Web-Site: http://www.ntis.gov/search/product.aspx?ABBR=ADA262672

Friedman, N. P. & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology, 133*(1), 101–135

Friston, K. J., Holmes, A. P., & Worsley, K. J. (1999). How many subjects constitute a study? *NeuroImage, 10*(1), 1-5.

Gonzalez, C., & Wimisberg, J. (2007). Situation awareness in dynamic decision making: Effects of practice and working memory. *Journal of Cognitive Engineering and Decision Making, 1*(1), 56-74.

Gopher, D. (1982). A selective attention test as a predictor of success in flight training. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 24*(2), 173-183.

Gugerty, L. (2011). Situation awareness in driving. In D. Fisher, M. Rizzo, M. K. Caird, & J. D. Lee, (Eds.), *Handbook of Driving Simulation for Engineering, 101 Medicine and Psychology*. Boca Raton, FL: Taylor and Francis.

Hart, S. G., Battiste, V., & Lester, P. T. (1984). *Popcorn: A supervisory control simulation for workload and performance research.* Paper presented at the 20th Annual Conference of Manual Control, NASA, Washington, DC.

Hirshfield, L. M., Girouard, A., Solovey, E. T., Jacob, R. J. K., Sassaroli, A., Tong, Y., & Fantini, S. (2007). *Human-computer interaction and brain measurement using functional near-infrared spectroscopy*. Poster session presented at the 20th ACM Symposium on User Interface Software and Technology, NewPort, RI.

Izzetoglu, M., Bunce, S., Izzetoglu, K., Onaral, B., & Pourrezaei, K. (2007). Function brain imaging using near-infrared technology - assessing cognitive activity in real-life situations. *IEEE Engineering in Medicine and Biology Magazine, 26*(4), 38-46.

Izzetoglu, M., Izzetoglu, K., Bunce, S., Ayaz, H., Devaraj, A., Onaral, B., & Pourrezaei, K. (2005). Functional near-infrared neuroimaging. *IEEE Transactions on Neural Systems and Rehabilitation Engineering,13*(2), 153-159.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences, 105*(19), 6829-6833.

James, W. (1952). *The principles of* psychology. Chicago, EUA : Encyclopaedia Britannica.

Jobsis, F. (1977). Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science, 198*(4323), 1264-1267.

Johannsdottir, K. R. (2004*). Situation awareness and working memory: An integration of an applied concept with a cognitive fundamental process.* Unpublished doctoral dissertation, Carleton University, Ottowa.

Johnston, W. & Heinz, S. (1978). Flexibility and capacity demands of attention. *Journal of Experimental Psychology, General, 107,*420-435.

Jones, G. V., & Martin, M. (2003). Individual differences in failing to save everyday computing work. *Applied Cognitive Psychology, 17*(7), 861-868.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to stroop interference. *Journal of Experimental Psychology: General, 132*(1), 47-70.

Kane, M. J., Poole, B. J., Tuholski, S. W., & Engle, R. W. (2006). Working memory capacity and the top-down control of visual search: Exploring the boundaries of "executive attention". *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(4), 749-777.

LaBerge, D. & Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review, 96,* 101-124.

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General, 133*(3), 339-354.

Macleod, C. (1991). Half a century of research on the stroop effect: An integrative review. *Psychological Bulletin, 109*, 163-203.

McCarley, J. S., Wickens, C. D., Goh, J., & Horrey, W. J. (2002). A computational model of Attention/Situation awareness. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 46*(17), 1669-1673.

Miller, G. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

Miller, E. K., Freedman, D. J., & Wallis, J. D. (2002). The prefrontal cortex: Categories, concepts and cognition. *Philosophical Transactions of the Royal Society of London.Series B: Biological Sciences,357*(1424), 1123-1136.

Osaka, M., Komori, M., Morishita, M., & Osaka, N. (2007). Neural bases of focusing attention in working memory: An fMRI study based on group differences. *Cognitive, Affective, & Behavioral Neuroscience, 7*(2), 130-139.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium*. Lawrence Erlbaum.

Salmon, P., Stanton, N., Walker, G., & Green, D. (2006). Situation awareness measurement: A review of applicability for C4i environments. *Applied Ergonomics, 37*(2), 225-238.

Sarter, N. B., Mumaw, R. J., & Wickens, C. D. (2007). Pilots' monitoring strategies and performance on automated flight decks: An empirical study combining behavioral and eye-tracking data. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 49*(3), 347-357.

Sarter, N. B., & Woods, D. D. (1991). Situation awareness: A critical but ill-defined phenomenon. *The International Journal of Aviation Psychology, 1*(1), 45-57.

Sassaroli, A., Zheng, F., Hirshfield, L. M., Girouard, A., Solovey, E. T., & Jacob, R. J. K. (2010). Discrimination of mental workload levels in human subjects with functional near-infrared spectroscopy. *Journal of Innovative Optical Health Sciences, 1*(2), 227-237.

Schneider, W. & R. M. Shiffrin. (1977). Controlled and automatic human information processing: 1. Detection, search, and attention. *Psychological Review*, *84,* 1-66.

Smith, K., & Hancock, P. A. (1995). Situation awareness is adaptive, externally directed consciousness. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37*, 137-148.

Sohn, Y. W., & Doane, S. M. (2004). Memory processes of flight situation awareness: Interactive roles of working memory capacity, long-term working memory, and expertise. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 46*(3), 461-475.

Sukthankar, R. (1997). *Situation awareness for tactical driving*. Unpublished doctoral dissertation, Carnegie Mellon University, Pittsburgh.

Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology, 12*(4), 242-248.

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

Tvaryanas, A. P., & Thompson, W. T. (2008). Recurrent error pathways in HFACS data: Analysis of 95 mishaps with remotely piloted aircraft. *Aviation, Space, and Environmental Medicine, 79*(5), 525-532.

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

Unsworth, N., Schrock, J. C., & Engle, R. W. (2004). Working memory capacity and the antisaccade task: Individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(6), 1302-1321.

Vaitkunas-Kalita, S., Landry, S. J., & Yoo, H. (2011). Coincidence between the scientific and folk uses of the term situation(al) awareness in aviation incident reports. *Journal of Cognitive Engineering and Decision Making, 5*(4), 378-400.

Villringer, A., & Chance, B. (1997). Non-invasive optical spectroscopy and imaging of human brain function. *Trends in Neurosciences, 20*(10), 435-442.

Villringer, A., Planck, J., Hock, C., Schleinkofer, L., & Dirnagl, U. (1993). Near infrared spectroscopy (NIRS): A new tool to study hemodynamic changes during activation of brain function in human adults. *Neuroscience Letters, 154*, 101-104.

Wickens, C. D. (1999). Cognitive factors in aviation. In F. T. Durso, R. S. Nickerson, R. W. Schvaneveldt, S. T. Dumais, D. S. Lindsay, & M. T. H. Chi (Eds.), *Handbook of Applied Cognition* (pp. 247–282). Wiley.

APPENDICES

APPENDIX A: DETAILED FLIGHT SCENARIO

For this study, flight scenario composed by Çak (2011) has been used with modifications for simulator pilots. These changes were made in cooperation with SMEs and they were mainly due to inadequacy of simulator pilots on some procedures of real piloting. Detailed scenario and SA queries will be described in two sections.

A.1 Flight Scenario

Participants were expected to perform a flight from İstanbul Atatürk Airport (LTBA) to Ankara Esenboğa Airport (LTAC). The given flight route is LTBA, Beykoz(BKZ), Yalova(YAA) and LTAC. BKZ and YAA are the given easy-to-fly navigation points with radio transmitters. Below, details of the flight scenario including pilot and experimenter instructions will be given.

After take-off from LTBA Runway 05, the participants climb in the runway heading to reach the BKZ waypoint. In this waypoint, a turn to YAA waypoint is executed. After the turn, experimenter worsens the weather conditions so that the participant pilots have difficult conditions increasingly. First, weather conditions are set be cloudy to alter visibility and windy to make controlling of the plane more difficult. On this part, participants are expected to be more cautious about the worsening conditions as their altitude gets higher. Later, experimenter introduces moderate icing, which causes malfunctioning of the airspeed indicator in the plane. The airspeed indicator is one of the most important instruments for the flight and its failure entails emergency landing. Experimenter informs the participant about the icing condition and asks if he wants to terminate the flight. If the participant wants to continue, the experimenter interferes in the following minute and directs the flight for landing back to LTBA.

In order to land back to Ataturk airport, participants are instructed to fly to an intersection point preceding the approach pattern. The intersection point, namely ERMAN, is a destination which is hard to find since no navigational device shows its location directly. In order to find it, the participants have to develop a strategy using at least two navigation instruments. During a period of 10 minutes, the experimenter waits for the participant to find his own strategy. If the participants fail to direct the plane to ERMAN, required heading information is given by the experimenter for the sake of the rest of the scenario. On the way to this point, the visibility is diminished with the increase of the cloud density to complicate the navigation. One hold is requested at ERMAN that will be followed by the approach for landing.

During the holding maneuver, horizontal situation indicator (HSI) and radio magnetic indicator (RMI) are set to fail. Under these circumstances, participants have to follow the magnetic compass very carefully as it is the last and not that precise navigational aid working in the plane for heading information. After the holding maneuver, participants are directed to descent for landing by following instrument landing system (ILS) procedures. Performing an ILS approach, required horizontal and vertical position of the plane can be read over glideslope and localizer equipments which are still functioning. At this stage, experimenter sets also vertical speed indicator (VSI) to fail. Due to the density of the cloud and low ceiling, participants cannot have visual contact with the runway and they are expected to execute missed approach procedures. After they cancel landing and start to ascend, the experimenter interrupts the flight and change the position and the altitude of the aircraft for the beginning of the approach pattern. To enable landing this time, the cloud ceiling is set to 600 ft. During the last approach for landing, airspeed indicator is also set to fail. At the end, the participant lands the airplane at the second approach to the runway. Further details of the scenario can be found at the Çak's dissertation (2011).

Modifications on the Çak's scenario are listed as follows:

- The participants are asked to climb to the final altitude of 9000 ft after take-off skipping the intermediate flight level in the original scenario to simplify the instructions. This simplification does not change any important aspect of the scenario and it even makes easier for the simulator pilots as aimed in the low workload period of the flight.
- Instead of the question about the wind direction, the amount of the fuel at the beginning of the flight is asked. Since this study is carried with simulator pilots who does not have any real flight experience, SMEs suggested to change this query with another one requiring general attention on the aircraft situation. Wind direction is quite hard to determine during the flight especially on a PC simulation environment. The new query is about the fuel level, which is a very important flight parameter that the pilot has to keep in mind all through the flight.
- The query asking the entry method for holding at ERMAN has been cancelled since the simulator pilots generally practice only one entry method for holds. SMEs reported that this question would not be distinguishing for simulator pilots. Instead of this question, participants are asked the position of the closest runway in clock direction. This is also a very important piece of information which the pilot has to recall in case of an emergency.
- Some timing instructions have been modified due to the differences in the simulation platforms between Çak's study and this one.
- The question "What will your entry method be for the holding pattern?" is changed with the question "What is the direction of the closest airport?". The original question contains a procedure which simulator pilots do not practice. SMEs suggested questioning important information that the pilots have to keep in mind during the whole flight.

The event list for the scenario is the same from Çak's study except for modifications. For the purpose of completeness, all events in this study are given below.

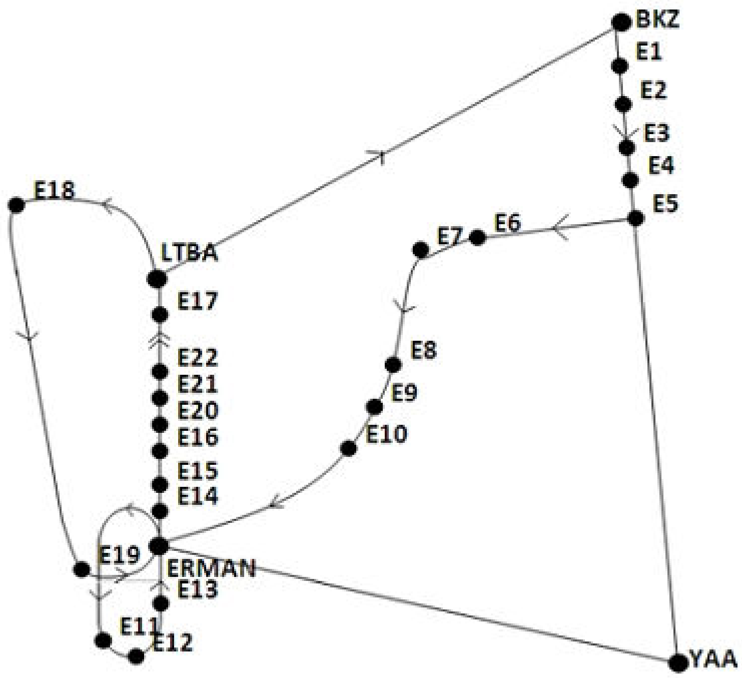


Figure A.1: Flight Route and Event Points in the Scenario (adopted from Çak (2011))

E1: Experimenter introduces clouds with overcast coverage.

E2: Experimenter introduces 20 knots of wind from the direction of 30° with moderate turbulence.

E3: Experimenter administers the query “"Where is the take-off airport in the clock direction?"

E4: Experimenter introduces moderate icing conditions.

E5: Experimenter directs the pilot to terminate the flight.

E6: Experimenter administers the query “What is your distance to ERMAN?”

E7: Pilot starts to follow the expected route to ERMAN.

E8: Experimenter administers the query "What is the amount of the fuel in the beginning of the flight?"

E9: Experimenter pauses the simulation and administers the first set of offline queries. (Appendix I)

E10: Experimenter deteriorates the visibility conditions to 0.8 km (low visibility).

E11: Experimenter freezes HIS and RMI gauges.

E12: Experimenter administers the query "What is the direction of the closest airport?"

E13: Experimenter administers the query "What is your heading?"

E14: Experimenter freezes the VSI gauge.

E15: Experimenter pauses the simulation and administers the second set of offline queries. (Appendix J)

E16: Experimenter administers the query "What is the missed approach procedure?"

E17-E18: Experimenter changes the position and altitude of the aircraft for the second approach.

E19: Experimenter increases the cloud ceiling to enable landing.

E20: Experimenter administers the query "What is your altitude?"

E21: Experimenter freezes the ASI gauge.

E22: Experimenter administers the query "What is your descent rate?"

A.2 SA Queries

In this section, accuracy criteria for SA queries will be described. An initial version for this assessment has been formed by the experimenter and examined by SMEs and Serkan Çak. Afterwards, the final version of the accuracy assessments has been composed by considering all comments from SMEs and Çak. Further details about queries can be found in Çak (2011).

Online SA Queries

- SAQ1. Where is the take-off airport in terms of clock direction?

The correct answer is accepted within ± 1 hour with regard to comments from SMEs and Çak. Partial credit (25%) is given for the answers in ± 2 hour range.

- SAQ2. What is your distance to ERMAN?

Accuracy for distance answer is calculated by the percentage formula given below. Accuracy (%) = 100 - [(Absolute value (answer - correct) / correct] *100

- SAQ3. What is the amount of fuel at the beginning?

Since the fuel information has been read over an analog gauge, SMEs suggested using a wide range of answers in evaluation. Correct answer received 100%, faulty with ± 5% received 75% and faulty with ± 10% got 25% credits.

- SAQ4. What is the clock direction to the closest runway?

Similar to the direction question in SAQ1, correct answer ± 1 hour got 100% and correct answer ± 2 hours got 25%.

- SAQ5. What is your heading?

According to similar SME comments, the correct answer ± 5 degrees got 100% and answers with 10 degrees of fault got 50% credits.

- SAQ6. What is the Missed Approach Procedure?

Since this question requires following a procedure completely, only totally correct answers received credits.

- SAQ7. What is your altitude?

For this query, a wide range for the answers is accepted. Correct answer ± 50 ft received full credits, answers with 100 ft faulty got 75% and answers with 200 ft faulty got 25%.

- SAQ8. What is your descent rate?

Based on the uniform agreement of the SMEs, correct answer is accepted with ± 100 ft/min. Answers with 200 ft/min faulty received 75% and those with 300 ft/min faulty got 25% credits.

Offline SA Queries

- SAOffline1. What is the indicated airspeed?

With respect to comments from SMEs, correct answers ± 5% are accepted and answers in ± 10% range got 25% credits.

- SAOffline2. What is your altitude in Mean Sea Level (MSL)?

Combining different comments from SMEs and Çak, correct answers ± 5% are accepted and answers in ± 10% range have been evaluated with 25% credits.

- SAOffline3. What is your heading?

Same accuracy criterion is used with SAQ5. The correct answer ± 5 degrees got 100% and ± 10 degrees fault range got 50% credits.

- SAOffline4. How much time will it take to the next control point?

Since this question is a hard one, even for the experimenter who can see the positions on the map, five different accuracy levels are used. Correct answers within ± 10% range got full points. Answers within ± 20% range got 75%, answers within ± 30% range got 50%, answers within ± 40% range got 25% and lastly, answers within ± 50% range got 10% credits.

- SAOffline5. What is your position on the map?

This query is evaluated by the experimenter with respect to the positions in the current status of the experiment.

- SAOffline6. What will be the next ATC call?

This query is required to be answered totally correctly since it is a block of information to be remembered.

- SAOffline7. What will be the first altitude to climb in the Missed Approach Procedure?

This query is also expected to be correctly answered due to the procedural information.

- SAOffline8. What is your ground speed?

The accuracy for this question is determined by the percentage calculation given as: Accuracy (%) = 100 - [(Absolute value (answer - correct)) / correct] *100

- SAOffline9. What is the approach ILS frequency?

This procedural information requires the only correct answer.

- SAOffline10. What is the decision altitude (DA) for the landing airport?

This procedural information requires the only correct answer.

- SAOffline11. What is the current altitude?

Similar to SAQ7, correct answer ± 50 ft received full credits, answers within ± 100 ft range got 75% and answers within ± 200 range got 25% credits.

- SAOffline12. What is the airport elevation?

Since this question requires a specific bit of information, only the correct answers are accepted.

- SAOffline13. What is the value of QNH?

For this query, the only correct answer is accepted.

A.3 Pre-Flight Scenario

Participants were asked to complete and record a pre-flight at home before the experiment. This pre-flight has been compiled in order to familiarize the participants for the novel events and special occurrences in the scenario of the experiment. The conditions in the actual experiment are not duplicated but similar situations have been generated.

The pre-flight starts and ends at Ankara Esenboğa Airport unlike the experiment. No equipment like maps, GPS or autopilot that assists navigation is allowed. Only flight charts for Ankara Esenboğa airport are allowed as in the experiment conditions. Further settings for the pre-flight are given below:

- Select Cessna C172SP as your airplane
- Set the Weather Theme as Fogged In
- Change fuel amount to 50% for Left land right storages
- Set equipment failures for Heading Indicator, ADF radio and Vertical Speed Indicator
- Open magnetic compass in the beginning of the flight.

Flight instructions are given as:

- Take off from Esenboğa Airport 03R runway, proceed to BUK VOR directly, and follow 132 outbound to KUBER Intersection point
- Keep your flight altitude at 7000 ft
- Perform 2 hold maneuvers at KUBER
- Direct your plane to Ankara NDB and land back to Esenboğa Airport using only Glideslope and Localizer in

APPENDIX B: PARTICIPANTS' BIOGRAPHICAL DATA

Table B.1: Participants' Biographical Data

Participant	Age	Total flight hours	Participation in fNIRS Study
1	36	970	*
2	49	5000	*
3	22	700	*
4	25	1100	*
5	30	1050	*
6	21	700	*
7	26	1250	*
8	40	800	
9	21	300	*
10	66	1000	*
11	30	200	
12	44	250	
13	59	130	
14	27	2000	
15	21	4500	
16	18	500	
17	46	8000	
18	18	370	
19	23	1350	
20	17	1000	
21	20	1300	
22	22	200	
23	18	550	
24	32	200	
25	64	13	
26	27	30	
27	25	1000	
28	24	50	
29	18	20	
30	45	10000	
31	31	10	
32	29	30	
33	37	500	
34	27	400	
35	18	2000	*

Table B.2: Participants' Biographical Data from Çak (2011) Study

Participant	Age	Total Flight Hours
1	31	2900
2	32	3500
3	32	2700
4	30	1600
5	46	506
6	46	4500
7	32	200
8	27	190
9	28	205
10	60	10000
11	33	1754
12	49	650
13	32	205
14	31	1000
15	30	203
16	30	205
17	31	1100
18	41	4200
19	31	2100
20	43	4200
21	37	1950
22	43	4675
23	27	850
24	37	2598
25	31	1250
26	32	675
27	27	255
28	30	320
29	21	610
30	26	514
31	35	1935
32	39	5100
33	36	3700
34	33	4530
35	34	2150
36	35	2440

APPENDIX C: TEST INSTRUCTIONS FOR THE EXPERIMENTER (In Turkish)

DENEY YÜRÜTÜCÜSÜ İÇİN YÖNERGE

Bu yönerge, deney yürütücüsünün uçuş yöneticisi ve ATC kule olarak görevlerini içerir.

Altı çizili olarak yazılmış cümleler ATC kule talimatlarını, >>> ile başlayan satırlar online durumsal farkındalık sorularını gösterir.

LTBA 05 Pist başında uçuş başlatıldığında pilot'a bu uçuşta;

T(ango)C(harlie)-S(ierra)A(lfa)W(hiskey) çağrı kodunu kullanacağı,

Autopilot kullanılmayacağı,

Uçağın Anti-Ice özelliğinin faal olarak simüle edilmiş olsa da senaryo gereği gayri-faal olduğu,

Shift-Z kısayoluyla açılan göstergelerin kullanılmayacağı,

Simulator ekranından herhangi bir harita açılmayacağı,

Pilot uçuş başlamadan önce pusulayı açması (shift + 5) gerektiği söylenir.

Yeşilköy kule: TC-SAW (Tango Charlie Sierra Alpha Whiskey), 05 pisti kalkış serbest, QNH 1013, doğrudan final seviyeniz 090.

Geri okuma onaylandıktan sonra pilotun 090'a tırmanması beklenir.

E1: BKZ üzerinden YAA VOR'una dönüş gerçekleşince.

Weather-Advanced Weather-Clouds

Add new Cloud Layer
Base: 300
Tops (MSL): 10000
Turbulence: Light
Cloud Coverage: Overcast

E2: Hemen sonrasında

Weather - Advanced Weather - Winds Surface – Winds Altitude: 10000
Speed (Knots): 20
Wind Direction: 30
Turbulence: Moderate

>>> E3: 1 dakika içinde
>>> Online Query1: Sormadan not ediniz – Kalkış havaalanı yönü:
>>> "Saat yönü cinsinden kalkış yapılan havaalanı neredeki?"

E4: Cevap alındıktan sonra

Weather - Advanced Weather – Clouds Icing: Moderate

Buzlanma pilot tarafından fark edilmezse 2 dakika sonra ATC tarafından bildirilir.

Yeşilköy yaklaşma: TC-SAW, orta seviyede buzlanma vardır. Uçuşunuza bu şartlarda devam etmek istiyor musunuz?

1 dakika içinde pilot uçuşu iptal etmezse;

Yeşilköy yaklaşma: TC-SAW, uçuşunuz hava muhalefeti nedeniyle iptal edilmiştir. ERMAN üzerinde 3000 feet alçalma serbest. ERMAN'da beklemeye alınacaksınız. Senaryo gereği bir süre boyunca ERMAN noktasına kendi çabanızla ulaşmanız beklenecektir.

E5: Pilotun bulunduğu noktadan IST VOR'una yönelerek ERMAN için 10 nm yarıçaplı bir ARC çizmesi ya da YAA üzerinden STAR Chart'ı okunarak SADIK ve daha sonra ERMAN ulaşması beklenir..

>>> E6: ERMAN için BKZ-YAA yolu terk edildikten sonra
>>> Online Query2: Instructor paneldeki radar görüntüsünü Print Screen tuşu ile alınız.
>>> "ERMAN noktasına olan uzaklığınız nedir?"

E7: Pilot'a ERMAN'ı bulması için 10 dk süre verilir.

Bu süre sonunda pilot'a vektör verilerek ERMAN'a ulaşması sağlanır.

>>> E8: Pilot uçağı Erman’a yönelttikten sonra

>>> Online Query3:

>>> “Başlangıçtaki yakıt miktarınız nedir?”

>>> E9: Hemen sonrasında

>>> Simülatör durdurularak katılımcıya Offline-1 soruları yöneltilir.

>>> IAS___ MSL___ Heading_____ ERMAN’a süre_____

E10: ERMAN’a yaklaşıldığında

Weather Visibility:1/2mi-0.8km

Weather-Advanced Weather-Clouds Base: 0

Yeşilköy yaklaşma: TC-SAW, ERMAN üzerinde bulunduğunuz uçuş başında 1 hold atın.

E11: Hold başlangıcında ilk bacakta

Failures-Instrument Heading Indicator: Failed ile HSI devre dışı bırakılır.

Failures-Radios ADF: Failed ile RMI devre dışı bırakılır.

>>> E12: Hold ‘un ilk dönüşünde

>>> Online Query4: Sormadan not ediniz – En yakın pist yönü:

>>> “En yakın pist yönü nedir?”

>>> E13: Hold süresince 2. Dönüşte

>>> Online Query: Sormadan not ediniz - Uçuş başı:

>>> “Uçuş başınız nedir?”

Yeşilköy yaklaşma: TC-SAW Görüş 830m Bulut Alttavanı 300ft’tir 35 Sağ için ILS yaklaşma olacak. Malumat.

E14: Hemen sonrasında

Failures-Instruments   Vertical Speed Indicator: Failed

>>> E15: Pilot alçalmaya başladığında

>>> Simülatör durdurularak Offline-2 soruları katılımcıya verilir.

>>> GS ____ Bir sonraki ATC mesajı __________ MSL ________

>>> E16: Online Query: Katılımcı ILS zarfına oturduktan sonra

>>> Pas geçme prosedürünüz nedir?”

E17-E18: Pas geçildikten sonra katılımcıya bilgi verilerek uçak Erman üzerine alınır. Katılımcı bilgisayarı üzerinde World – Map – Uçak mouse ile taşınır, uçuş başı 355, irtifa 3000 ayarlanır.

E19: Hemen sonrasında

Weather-Advanced Weather-Clouds   Base: 600ft

Yeşilköy Kule: 35R iniş serbest. Rüzgar 030dan 20 knot. QNH 1013.

>>> E20: Alçalma devam ederken,

>>> Online Query:Sormadan not ediniz – İrtifa:

>>> “İrtifanız nedir?”

E21: Hemen sonrasında,

Failure-Instruments   Airspeed Indicator: Failed

>>> E22: ILS zarfına girildiğinde

>>> Online Query: Sormadan not ediniz – vertical speed:

>>> “Alçalma varyonuz nedir?”

APPENDIX D: INFORMED CONSENT FORM
(In Turkish)

Gönülü Katılım Formu

Bu çalışma, ODTÜ Enformatik Enstitüsü Bilişsel Bilimler Ana Bilim Dalı Yüksek Lisans öğrencisi Orçun Orkan ÖZCAN, ve tez danışmanları Dr. Murat Perit ÇAKIR ve Dr. Bilge SAY tarafından durumsal farkındalığın çalışan bellek ve dikkat mekanizmalarını ortaya çıkarmaya yönelik bir çalışmadır. Bu kapsamda Windows Flight Simulator uçuş simülasyonunda deneyimli oyuncular yani sanal pilotlardan bilgi toplamak 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 kimlik bilgileri herhangi bir şekilde kullanılmadan sadece bilimsel yayınlarda kullanılacaktır.

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 söylemek yeterli olacaktır. Deney sonunda, bu çalışmayla ilgili sorularınız cevaplanacaktır. Bu çalışmaya katıldığınız için şimdiden teşekkür ederiz. Çalışma hakkında daha fazla bilgi almak için Bilişsel Bilimler Bölümü öğretim üyelerinden Dr. Murat Perit ÇAKIR (Tel: 210 7706; E-posta: perit@ii.metu.edu.tr), Dr. Bilge SAY (Tel: 210 7139; E-posta: bsay@ii.metu.edu.tr) ya da yüksek lisans öğrencisi Orçun Orkan ÖZCAN (Tel: 210 7720; E-posta: orkan@ii.metu.edu.tr) 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	Tarih	İmza
	----/----/-----	

APPENDIX E: TEST INSTRUCTIONS FOR PARTICIPANTS (In Turkish)

DURUMSAL FARKINDALIK ÖLÇÜM TESTİNE
KATILACAKLAR İÇİN YÖNERGE

Deney için bilinmesi gerekenler:

TC-SAW çağrı kodunu kullanacak,
Autopilot kullanılmayacak,
Uçağın Anti-Ice özelliğinin kullanmayacak,
Shift-Z kısayoluyla açılan göstergelerin kullanılmayacak,
Simulator ekranından herhangi bir harita (GPS dahil) açılmayacaktır.
Uçuşa başlamadan önce pusula (shift+5) açılmalıdır.

1. Atatürk Havalimanından (LTBA), Esenboğa Havalimanına (LTAC) bir uçuş gerçekleştirmeniz beklenmektedir.

2. Uçuş rotanız LTBA DCT BKZ DCT YAA DCT LTAC olacaktır. İrtifanız 9,000ft'tir. Yalova'dan sonra Esenboğa'ya kadar direkt vektör verilecektir.

3. Atatürk Havalimanında görünürlük 800m, bulutalttavanı 600ft'tir. Meteorolojik şartların (buzlanma, yağmur, rüzgar vs.) kötüleşme ihtimali vardır.

4. Atatürk Havalimanında kalkış pisti olarak RWY05' i kullanmanız beklenmektedir.

5. Kalkış ve iniş pistlerinin SID/STAR kartlarının eğitmen tarafından size verildiğinden emin olunuz. Kartları incelemek için kendinize yeteri kadar zaman ayırınız.

6. Yeteri kadar brifing yaptığınızdan emin olduğunuzda eğitmene bildiriniz ve ATC'nin kleransına müteakip uçuşa başlayınız

7. Eğitmen, ATC, kule ve ATIS olarak uçuşunuz sırasında size gerekli bilgileri verecektir.

8. Uçuş sırasında size sorulan uçuşla ilgili bilgilere mümkün olduğunca çabuk, cümle kurmadan, soruyu tekrar etmeden net cevaplar vermeye çalışınız.

a. Örneğin "Uçuş irtifanız nedir?" sorusunu sadece irtifa bilgisini vererek eğitmen tarafından duyulacak şekilde cevaplayınız.

b. Yaptığınız hesaplamaları, ara adımları cevabınıza dahil etmemeye çalışınız

9. Simülatör, uçuşun belirli bölümlerinde durdurulacak ve sizden kağıt üzerinde yazılı sorulara uçuş göstergelerine ve kullandığınız uçuş kartlarına bakmadan yazılı olarak cevap vermeniz istenecektir. Hafızanızda kalan bigilere göre en yakın cevabı işaretlemeye çalışınız

10. Soruları cevaplamanız ve soru/cevap formunu araştırmacıya teslim etmenizin ardından uçuşunuz kaldığı yerden devam edecektir.

11. Konsantrasyonunuzun bozulmaması ve ölçümlerin sağlığı açısından uçuş sırasında cep telefonlarınızın kapalı olması önemlidir.

APPENDIX F: PILOT WORKLOAD SURVEY
(In Turkish)

PİLOT İŞ YÜKÜ DEĞERLENDİRME FORMU

Katılımcı No:

---

1. Lütfen uçuş sırasındaki **işyükü oranına** karşılık gelen şıkkı işaretleyiniz

0 1 2 3 4 5 6

Düşük Yüksek

2. Lütfen simülasyon uçuşunda **kendi performansınızı** değerlendiriniz

0 1 2 3 4 5 6

Çok Zayıf Çok İyi

3. Gerçekleştirdiğiniz uçuş için ne kadar **dikkat** ihtiyacı duydunuz?

0 1 2 3 4 5 6

Çok Az Çok Fazla

4. Uçuş ne kadar **karmaşıktı**?

0 1 2 3 4 5 6

Değildi Çok Fazla

5. Uçuş sırasında ne kadar **zaman baskısı** hissettiniz?

0 1 2 3 4 5 6

Hiç Çok Fazla

6. Uçuş hangi seviyede **zihinsel efor** gerektirdi?

0	1	2	3	4	5	6
Hiç						Çok Fazla

7. Uçuş sırasındaki **yoğunluğunuzun** seviyesi neydi?

0	1	2	3	4	5	6
Hiç						Çok Fazla

8. Uçuşun **zorluk derecesi** neydi?

0	1	2	3	4	5	6
Kolay						Zor

9. Lütfen bu uçuştaki **motivasyonunuzu** değerlendiriniz?

0	1	2	3	4	5	6
Düşük						Yüksek

10. Uçuştan sonra kendinizi nasıl **hissettiniz**?

0	1	2	3	4	5	6
Dinç						Yorgun

0	1	2	3	4	5	6
Sakin						Gergin

11. Aşağıdaki uçuş fazlarında maruz kaldığınız **işyükünü** notlayınız.

Kalkış

0	1	2	3	4	5	6
Düşük						Yüksek

LTBA-BKZ Arası

0	1	2	3	4	5	6
Düşük						Yüksek

BKZ-YAA Arası

0 1 2 3 4 5 6

Düşük Yüksek

BKZ-YAA'dan ERMAN'a dönüş

0 1 2 3 4 5 6

Düşük Yüksek

ERMAN'da bekleme

0 1 2 3 4 5 6

Düşük Yüksek

Yaklaşma ve Pas Geçme

0 1 2 3 4 5 6

Düşük Yüksek

APPENDIX G: DEBRIEFING SHEET
(In Turkish)

KATILIM SONRASI BİLGİ FORMU

Bu çalışma daha önce de belirtildiği gibi ODTÜ Bilişsel Bilimler Bölümü yüksek lisans öğrencisi Orçun Orkan ÖZCAN ve öğretim üyelerinden Dr. Murat Perit ÇAKIR ve Dr. Bilge SAY tarafından yürütülen durumsal farkındalık üzerine bir çalışmadır. Bu çalışmada temel olarak, deneyim seviyesi ve kişisel bilişsel farklılıkların durumsal farkındalığı nasıl etkilediği, durumsal farkındalıkla toplam uçuş saati, çalışma belleği ve dikkat mekanizmaları arası ilişki incelenecektir.

Durumsal Farkındalık, operasyonel olarak tanımlanmış ve birçok bilişsel yeteneği kapsayan bir kavramdır. En genel tanımı ile operatörün çevresini algılaması, bunu yorumlaması ve gelecek hakkında tahmin yapmasıdır. Durumsal Farkındalık konusundaki mevcut çalışmalar kavramın psikolojik temellerine dayanmamaktadır ve tanımlanması konusunda bile fikir ayrılıkları bulunmaktadır. Bu çalışmada, durumsal farkındalık kavramı yapılan deneylerle incelenecek ve deneyim ve bilişsel farklılıkların durumsal farkındalık ile ilişkileri araştırılacaktır.

Bu mekanizmaları ortaya çıkarmak amacıyla bu deneyde, çalışma belleği, ketleme ve bölünmüş dikkat üzerine psikolojik testler uygulanmıştır.

Bu çalışmadan alınacak ilk verilerin Mart 2012 sonunda elde edilmesi amaçlanmaktadır. Elde edilen bilgiler sadece bilimsel araştırma ve yazılarda kullanılacaktır. Çalışmanın sonuçlarını öğrenmek ya da bu araştırma hakkında daha fazla bilgi almak için aşağıdaki isimlere başvurabilirsiniz. Bu araştırmaya katıldığınız için tekrar çok teşekkür ederiz.

Dr. Murat Perit ÇAKIR (Tel: 210 7706; E-posta: perit@ii.metu.edu.tr)
Dr. Bilge SAY (Tel: 210 7139; E-posta: bsay@ii.metu.edu.tr)
Orçun Orkan ÖZCAN (Tel: 210 7720; E-posta: orkan@ii.metu.edu.tr)

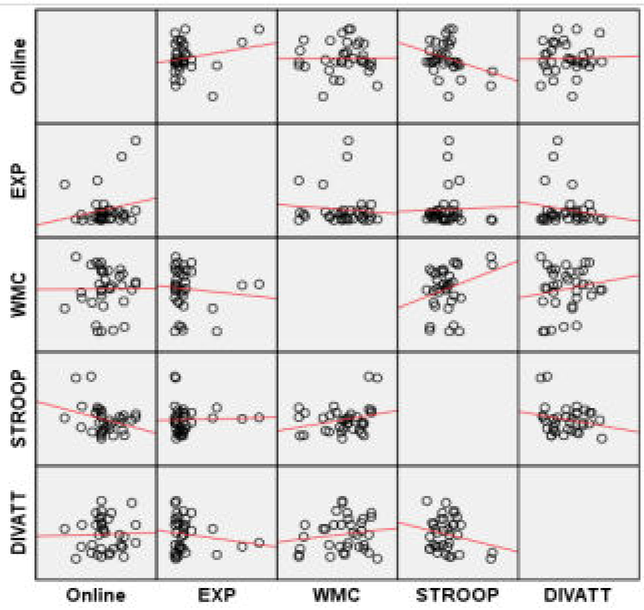
APPENDIX H: EXPERIMENT RESULTS

H.1 Descriptive Statistics for all variables

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Online	35	88	678	424,54	133,776	-,153	,398	,302	,778
OnlineRT	35	3,00	62,25	10,3816	10,01778	4,295	,398	22,046	,778
Offline	35	294	963	630,14	161,469	-,574	,398	,133	,778
Combined	35	526	1371	1054,69	229,407	-,885	,398	-,035	,778
EXP	35	10	10000	1356,37	2217,584	2,835	,398	8,089	,778
EXP_L	35	1,00	4,00	2,6603	,75199	-,572	,398	,003	,778
WMC	35	3	68	39,94	18,383	-,687	,398	-,287	,778
STROOP	35	-19,52	1068,93	310,1799	242,34337	1,558	,398	3,529	,778
DIVATT	35	2	31	14,23	8,257	,268	,398	-,974	,778
Valid N (listwise)	35								

H.2 Analysis for Online SA scores



Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	DIVATT, EXP_L, WMC, STROOP[b]	.	Enter

a. Dependent Variable: Online
b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,335[a]	,112	-,006	134,187

a. Predictors: (Constant), DIVATT, EXP_L, WMC, STROOP

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	68279,143	4	17069,786	,948	,450[b]
	Residual	540181,543	30	18006,051		
	Total	608460,686	34			

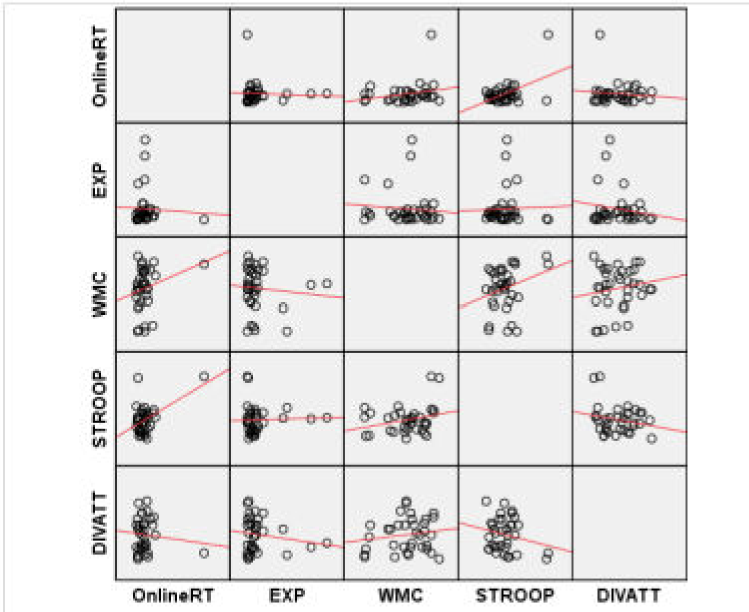
a. Dependent Variable: Online
b. Predictors: (Constant), DIVATT, EXP_L, WMC, STROOP

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	435,249	118,640		3,669	,001
	EXP_L	10,783	31,884	,061	,338	,738
	WMC	,834	1,337	,115	,624	,538
	STROOP	-,186	,106	-,338	-1,750	,090
	DIVATT	-1,046	2,943	-,065	-,355	,725

a. Dependent Variable: Online

H.3 Analysis for Online RT measurements



Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	DIVATT, EXP_L, WMC, STROOP[b]	.	Enter

a. Dependent Variable: OnlineRT
b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,518[a]	,269	,171	9,12006

a. Predictors: (Constant), DIVATT, EXP_L, WMC, STROOP

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	916,835	4	229,209	2,756	,046[b]
	Residual	2495,264	30	83,175		
	Total	3412,099	34			

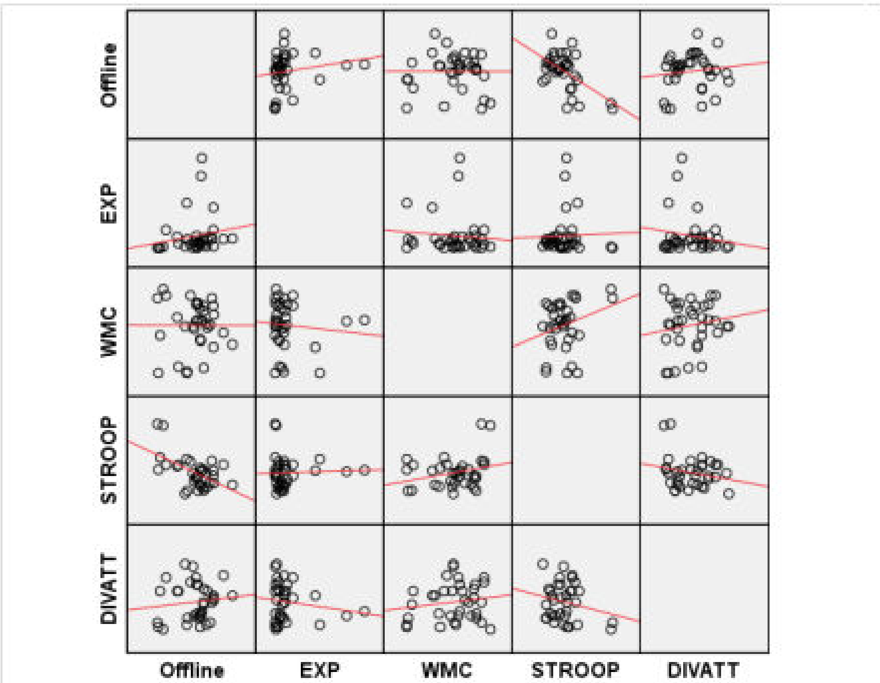
a. Dependent Variable: OnlineRT
b. Predictors: (Constant), DIVATT, EXP_L, WMC, STROOP

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	6,096	8,063		,756	,456
	EXP_L	-1,258	2,167	-,094	-,581	,566
	WMC	,064	,091	,118	,709	,484
	STROOP	,018	,007	,434	2,481	,019
	DIVATT	-,036	,200	-,030	-,179	,859

a. Dependent Variable: OnlineRT

H.4 Analysis for Offline SA scores



Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	DIVATT, EXP_L, WMC, STROOP[b]	.	Enter

a. Dependent Variable: Offline
b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,674[a]	,454	,381	127,034

a. Predictors: (Constant), DIVATT, EXP_L, WMC, STROOP

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	402325,251	4	100581,313	6,233	,001[b]
	Residual	484127,035	30	16137,568		
	Total	886452,286	34			

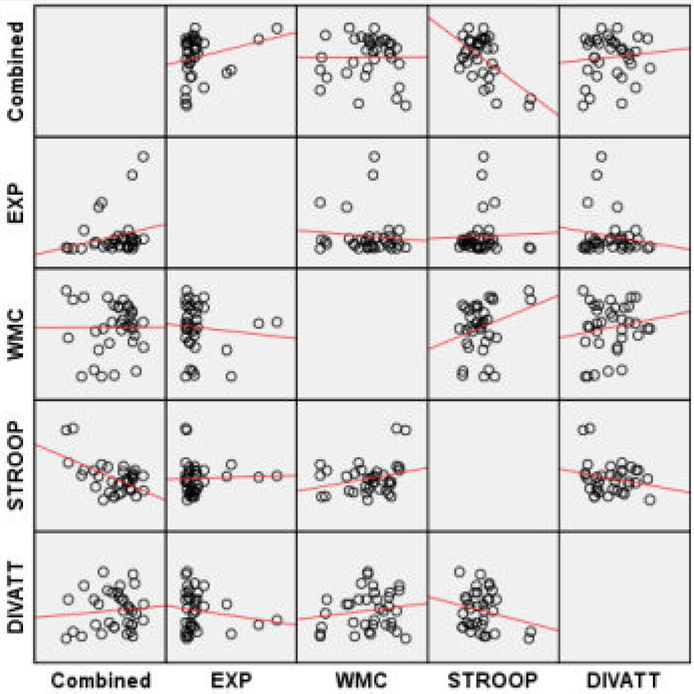
a. Dependent Variable: Offline
b. Predictors: (Constant), DIVATT, EXP_L, WMC, STROOP

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	464,189	112,316		4,133	,000
	EXP_L	79,951	30,185	,372	2,649	,013
	WMC	1,438	1,265	,164	1,137	,265
	STROOP	-,330	,101	-,495	-3,270	,003
	DIVATT	-,136	2,786	-,007	-,049	,961

a. Dependent Variable: Offline

H.5 Analysis for Combined SA scores



Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	DIVATT, EXP_L, WMC, STROOP[b]	.	Enter

a. Dependent Variable: Combined
b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,658[a]	,433	,358	183,870

a. Predictors: (Constant), DIVATT, EXP_L, WMC, STROOP

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	775098,779	4	193774,695	5,732	,002[b]
	Residual	1014242,764	30	33808,092		
	Total	1789341,543	34			

a. Dependent Variable: Combined
b. Predictors: (Constant), DIVATT, EXP_L, WMC, STROOP

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	899,438	162,567		5,533	,000
	EXP_L	90,734	43,690	,297	2,077	,046
	WMC	2,272	1,831	,182	1,241	,224
	STROOP	-,516	,146	-,545	-3,536	,001
	DIVATT	-1,182	4,033	-,043	-,293	,771

a. Dependent Variable: Combined

H.6 Analysis for Percieved Workload

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Low_WL	1,7571	35	1,31938	,22302
	High_WL	3,9786	35	1,04736	,17704

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Low_WL & High_WL	35	,480	,003

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Low_WL - High_WL	-2,22143	1,22890	,20772	-2,64357	-1,79929	-10,694	34	,000

H.7 Analysis for fNIRS measurements

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	V1_LowWL	1,8779	10	3,24204	1,02522
	V1_HighWL	19,3917	10	13,28786	4,20199
Pair 2	V2_LowWL	1,8466	9	3,91581	1,30527
	V2_HighWL	13,9133	9	21,09587	7,03196
Pair 3	V3_LowWL	2,3431	10	3,97272	1,25628
	V3_HighWL	26,3670	10	14,10911	4,46169
Pair 4	V4_LowWL	,9187	9	3,19866	1,06622
	V4_HighWL	13,4734	9	13,36475	4,45492
Pair 5	V5_LowWL	1,2084	9	2,40532	,80177
	V5_HighWL	21,3783	9	14,37875	4,79292
Pair 6	V6_LowWL	,9161	9	2,42966	,80989
	V6_HighWL	10,0100	9	12,92828	4,30943
Pair 7	V7_LowWL	1,2872	9	2,38371	,79457
	V7_HighWL	17,8491	9	12,59221	4,19740
Pair 8	V8_LowWL	,8397	9	2,14416	,71472
	V8_HighWL	4,9540	9	17,31682	5,77227
Pair 9	V9_LowWL	1,2474	9	2,18253	,72751
	V9_HighWL	18,2739	9	17,03597	5,67866
Pair 10	V10_LowWL	1,1624	6	1,53555	,62688
	V10_HighWL	10,9019	6	13,58112	5,54447
Pair 11	V12_LowWL	1,7777	10	2,20948	,69870
	V12_HighWL	15,4545	10	14,74815	4,66377
Pair 12	V14_LowWL	1,7955	9	2,95432	,98477
	V14_HighWL	18,8547	9	13,90966	4,63655
Pair 13	V15_LowWL	1,0700	9	1,40145	,46715
	V15_HighWL	19,4214	9	10,69811	3,56604
Pair 14	V16_LowWL	1,4771	9	3,16685	1,05562
	V16_HighWL	15,8879	9	15,71919	5,23973

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	V1_LowWL & V1_HighWL	10	,800	,005
Pair 2	V2_LowWL & V2_HighWL	9	,945	,000
Pair 3	V3_LowWL & V3_HighWL	10	,731	,016
Pair 4	V4_LowWL & V4_HighWL	9	,945	,000
Pair 5	V5_LowWL & V5_HighWL	9	,779	,013
Pair 6	V6_LowWL & V6_HighWL	9	,740	,023
Pair 7	V7_LowWL & V7_HighWL	9	,892	,001
Pair 8	V8_LowWL & V8_HighWL	9	,855	,003
Pair 9	V9_LowWL & V9_HighWL	9	,807	,009
Pair 10	V10_LowWL & V10_HighWL	6	,860	,028
Pair 11	V12_LowWL & V12_HighWL	10	,660	,038
Pair 12	V14_LowWL & V14_HighWL	9	,863	,003
Pair 13	V15_LowWL & V15_HighWL	9	,518	,154
Pair 14	V16_LowWL & V16_HighWL	9	,938	,000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	V1_LowWL - V1_HighWL	-17,51381	10,86822	3,43683	-25,28847	-9,73915	-5,096	9	,001
Pair 2	V2_LowWL - V2_HighWL	-12,06671	17,44079	5,81360	-25,47288	1,33947	-2,076	8	,072
Pair 3	V3_LowWL - V3_HighWL	-24,02389	11,52923	3,64586	-32,27141	-15,77638	-6,589	9	,000
Pair 4	V4_LowWL - V4_HighWL	-12,55471	10,39619	3,46540	-20,54593	-4,56349	-3,623	8	,007
Pair 5	V5_LowWL - V5_HighWL	-20,16989	12,59467	4,19822	-29,85101	-10,48876	-4,804	8	,001
Pair 6	V6_LowWL - V6_HighWL	-9,09391	11,24848	3,74949	-17,74025	-,44756	-2,425	8	,042
Pair 7	V7_LowWL - V7_HighWL	-16,56196	10,52041	3,50680	-24,64866	-8,47526	-4,723	8	,001
Pair 8	V8_LowWL - V8_HighWL	-4,11432	15,52453	5,17484	-16,04753	7,81889	-,795	8	,450
Pair 9	V9_LowWL - V9_HighWL	-17,02650	15,33012	5,11004	-28,81027	-5,24272	-3,332	8	,010
Pair 10	V10_LowWL - V10_HighWL	-9,73950	12,28590	5,01570	-22,63276	3,15376	-1,942	5	,110
Pair 11	V12_LowWL - V12_HighWL	-13,67681	13,39212	4,23496	-23,25695	-4,09666	-3,230	9	,010
Pair 12	V14_LowWL - V14_HighWL	-17,05923	11,45842	3,81947	-25,86696	-8,25151	-4,466	8	,002
Pair 13	V15_LowWL - V15_HighWL	-18,35144	10,04459	3,34820	-26,07240	-10,63049	-5,481	8	,001
Pair 14	V16_LowWL - V16_HighWL	-14,41083	12,79518	4,26506	-24,24608	-4,57559	-3,379	8	,010

H.8. Analysis for Comparison with Çak's Study

Group 0: Professional Pilots from Çak's study

Group 1: Simulator Pilots from the current study

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
OnlineRT	0	35	13,1205	8,63616	1,45978
	1	35	10,0595	18,15904	3,06944
Offline	0	35	720,37	195,802	33,097
	1	35	602,86	166,565	28,155
EXP	0	36	2017,58	2085,340	347,557
	1	35	1356,37	2217,584	374,840
WMC	0	36	37,94	15,183	2,530
	1	35	39,94	18,383	3,107
STROOP	0	36	458,445	275,694	45,949
	1	35	369,668	285,191	48,206

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Online RT	Equal variances assumed	,197	,659	,901	68	,371	3,06097	3,39888	-3,72140	9,84333
	Equal variances not assumed			,901	48,632	,372	3,06097	3,39888	-3,77065	9,89258
Offline	Equal variances assumed	,696	,407	2,704	68	,009	117,514	43,452	30,807	204,221
	Equal variances not assumed			2,704	66,296	,009	117,514	43,452	30,767	204,262
EXP	Equal variances assumed	,379	,540	1,295	69	,200	661,212	510,728	-357,662	1680,086
	Equal variances not assumed			1,294	68,446	,200	661,212	511,176	-358,703	1681,127
WMC	Equal variances assumed	,923	,340	-,500	69	,619	-1,998	3,996	-9,971	5,974
	Equal variances not assumed			-,499	65,896	,620	-1,998	4,007	-9,999	6,003
STROOP	Equal variances assumed	,031	,861	1,334	69	,187	88,777	66,567	-44,016	221,570
	Equal variances not assumed			1,333	68,732	,187	88,777	66,597	-44,090	221,643

APPENDIX I: FORM FOR OFFLINE SA QUERIES (Set #1)
(In Turkish)

Eğitmen Bölümü

Katılımcı No :

Kalkış Zamanı :

Simülatör Durdurma Zamanı :

Katılımcı Bölümü

1. IAS (Indicated Air Speed) değeri tahmini olarak nedir?

----

2. Uçuş irtifanız (MSL) tahmini olarak nedir?

----

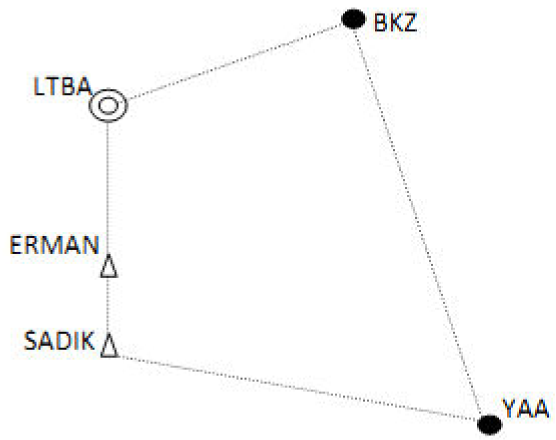
3. Uçuş başınız (heading) tahmini olarak nedir?

----

4. Bir sonraki kontrol noktasına tahmini olarak ne kadar süre sonra ulaşacaksınız?

----

5. Pozisyonunuzu LTBA, BKZ, YAA, SADIK ve ERMAN noktalarına göre artı (+) işaretiyle gösteriniz?


BKZ
LTBA
ERMAN
SADIK
YAA

APPENDIX J: FORM FOR OFFLINE SA QUERIES (Set #2)
(In Turkish)

Eğitmen Bölümü

Katılımcı No :

Kalkış Zamanı :

Simülatör Durdurma Zamanı :

Katılımcı Bölümü

1. Bir sonraki ATC mesajının en yüksek olasılıkla ne olmasını bekliyorsunuz?

----

2. Pas geçtikten sonra ilk tırmanacağınız irtifa ne olmalıdır?

----

3. Yer hızınız (GS) tahmini olarak nedir?

----

4. Yaklaşma ILS frekansı nedir?

----

5. İniş pisti için Karar İrtifası (DA) nedir?

----

6. Geçilen İrtifa nedir?

----

7. Meydan İrtifası nedir?

----

8. Mahalli Basınç (QNH) nedir?

ODTÜ

ENFORMATİK ENSTİTÜSÜ

YAZARIN

Soyadı : ÖZCAN

Adı : ORÇUN ORKAN

Bölümü : BİLİŞSEL BİLİMLER

TEZİN ADI (İngilizce) : EXPLORING THE EFFECTS OF WORKING MEMORY CAPACITY, ATTENTION, AND EXPERTISE ON SITUATION AWARENESS IN A FLIGHT SIMULATION ENVIRONMENT

TEZİN TÜRÜ : Yüksek Lisans ..X.. Doktora ....

1) Tezimden fotokopi yapılmasına izin vermiyorum ☐

2) Tezimden dipnot gösterilmek şartıyla bir bölümünün fotokopisi alınabilir ☐

3) Kaynak gösterilmak şartıyla tezimin tamamının fotokopisi alınabilir ☐

Yazarın imzası .... Tarih ....