INVESTIGATING CYBER PERFORMANCE:AN INDIVIDUAL DIFFERENCES STUDY

by

Kelly A. Cole

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

Purdue Polytechnic Institute of Technology

West Lafayette, Indiana

August 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Marcus Rogers, Chair

Department of Computer and Information Technology

Dr. Alexander Francis, Co-Chair

Department of Speech, Language and Hearing Sciences

Dr. Baijian Yang

Department of Computer and Information Technology

Dr. Umit Karabiyik

Department of Computer and Information Technology

Approved by:

Dr. Kathryne Newton

To my three sons: Morrison Stephen Cole Dale, Bennett Richard Cole Dale and Griffin Woods Cole Dale

"He says the best way out is always through. And I agree to that, or in so far as that I can see no way out but through" (Frost, 1915, p. 64).

ACKNOWLEDGMENTS

Thank you to Dr. Marcus Rogers for believing in me, even when I did not believe in me! Although I am forever grateful to Dr. Rogers for giving me this opportunity, I am also grateful for his wife June Rogers for her unwavering support for me and for all of the graduate students that enter Dr. Roger’s Cyber Forensics program. I would like to thank Dr. Alexander Francis for his invaluable guidance and without whom this would not have been possible. I am grateful for all of the support that I received from Dr. Baijian Yang and Dr. Umit Karabiyik. My committee members could see the entire forest, while at many times I could only see the trees; without them I may have never found my way through a PhD. I want to further thank Joseph Balazs for his data collection efforts and friendship. I am appreciative of the support I received from Dr. Chris James Forsythe of Sandia National Laboratory, Dr. Jeremy Wolfe, Dr. Justin Kantner (Boiler-Up), Dr. Joshua Koen, among the many others whom took the time to meet with me and/or to answer all of my questions. Last but not least, I want to recognize Kaethe Beck as another guiding light on the path towards my PhD degree, and for everything that she did to help me achieve it. I am thankful for the DHS scholarship and Lockheed Martin scholarship that I was awarded to support my research efforts. Lastly, I want to thank my Mom (GeGe) for all of the times she selflessly watched the kids for me so that I could write the following dissertation in peace.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

GLOSSARY

Behavior – What people do (Motowildo, Borman, & Schmit, 1997).

ICMP Attack – Internet control message protocol attacks (ICMP) is the network layer concerned with Internet services such as routing. Network layer DoS attacks are used to consume the network bandwidth and they can include ICMP flood attacks and user datagram protocol (UDP) floods (Compton & Hornat, 2007; Das, Karabade, & Tuna, 2015).

ICMP Flood – For a ICMP flood attack, a large volume of ICMP ECHO REQUEST packets (e.g., ping requests) are sent to overload the targeted network until it can no longer handle legitimate traffic. That results in legitimate users being unable to access the network (Compton & Hornat, 2007).

Incident Detection System – The IDS includes a database of known attack signatures or event patterns. It will search the signatures in packets and network traffic and if there is a match, it issues an alert (Shanks, 2015).

TCP Attacks – TCP attacks are concerned with the reliability of data communications between client and server. The attacks affect any computer providing TCP-based network services such as a Web server or a FTP server (Compton & Hornat, 2007).

WIZ Attack – A form of distributed denial of service (DDoS) attacks that are designed to cause Internet resources (i.e. websites, servers, routers) to become overwhelmed and then unavailable to users (Das et al., 2015). Expensive DDoS attacks were found to be the most common network attacks targeting enterprises in Cisco’s 2014 Annual Security Report (Cisco, 2014).

Work Performance – The expected organizational value of what people do (Motowildo et al., 1997).

ABSTRACT

The persistent issues that have been identified in the cyber defense domain, such as information-overload, burn-out and high turn-over rates among cyber analysts leads us to question what the cognitive ability contribution is to a more successful cyber performance. Cyber defense researchers theorize that individual differences are determinants of cyber performance success but have yet to establish empirically the role of individual differences. Therefore, the study uses an individual differences approach under a work performance framework to study the contributions of cognitive ability (i.e., attention control) on cyber performance success in a specific cyber work-role (i.e., the Incident Reponder), and through its well-defined primary task (i.e., incident detection system performance). The sample included actual network analysts with a wide range of incident detection expertise, age, and education levels for more reliable and valid scores. The results of the correlational analysis showed that individual differences in attention control (i.e., flexibility and spatial attention) contribute most to the differences in Incident Responder work-performance. A linear regression model then demonstrated that spatial attention and flexibility predict 53 to 60 percent of the variance in cyber performance scores. It is suggested that the KSA's from the NICE framework be updated with the cognitive abilities that contribute to and/or predict cyber performance success, for superior recruitment efforts towards a more efficient cyber defense work-force.

CHAPTER 1. INTRODUCTION

This chapter provides an overview of the research study. It introduces the research by presenting a background of the problem area and research questions. In addition, it covers the research significance, assumptions, limitations and delimitations which define the extent of the study.

1.1 Background

The identification of specific human factors (i.e., knowledge, skills and abilities) (Newhouse, Keith, Scribner, & Witte, 2017) that contribute to cyber performance is important for improving the efficiency of cyber security efforts. Unfortunately, the cyber analyst is very understudied in the human-factors research (R. S. Gutzwiller, Hunt, & Lange, 2016) and therefore, the cognitive ability contribution to cyber performance has not yet been identified. The current dissertation serves as a starting point for scientifically identifying the specific cognitive abilities that underlie cyber performance success. An individual differences framework (Cronbach, 1957; Hunter, 1983; Motowildo et al., 1997) that acknowledges the relationship between an individuals cognitive ability, knowledge and skills on task performance is applied in order to accomplish that goal. The dissertation showed that the current methods used to investigate the human factors contributing to cyber performance are inadequate and often not precise. The following chapters explain the importance of the individual differences research, offers an overview of previous research, provides the methodology used in the study and an analysis of the results.

1.2 Problem Statement

A major problem today for organizations connected to the Internet is to become a victim to cyber attack. That is because a successful attack results in heavy societal and monetary losses (Riley, Elgin, Lawrence, & Matlack, 2014). Despite declaring cyber defense efforts a national security mission (Mancuso, Strang, Funke, & Finomore, 2014) and the increasing technological advancements, cyber attacks continue to rise while the security operation centers (SOC) remain low performing (Cisco, 2020; Ponemon, 2020). According to a 2019 and 2020 survey by Ponemon Institute (2020) of over 600 information technology and security professionals, only 50 percent of them rated their SOC as effective in the ability to gather evidence, investigate and find the source of threats, although effectiveness did improve slightly from 2019 to 2020. The Figure in 1.1 displays the main issues identified in the SOC survey, such as an increasing work-load that leads to burn-out, lack of expertise and/or talent, high turn-over, information overload, etc. The CISO Benchmark Report (Cisco, 2020) based on a double-blind study of 2800 global participants conducted in late 2019 identify similar findings.

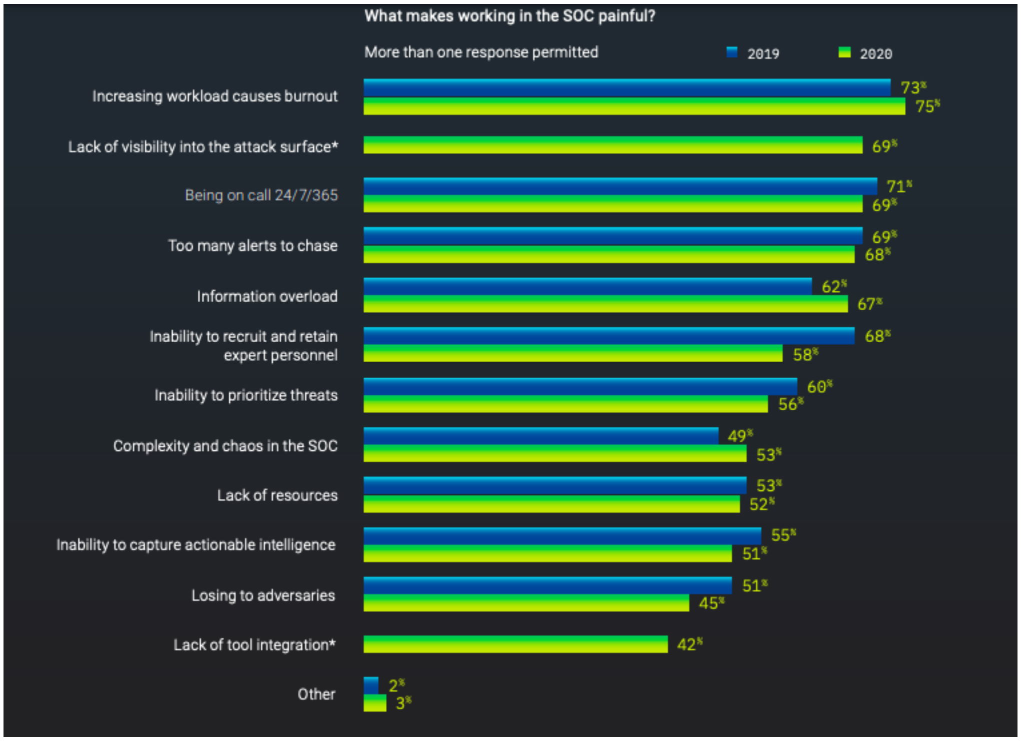


Figure 1.1. Main Issues Identified in the SOC (Ponemon, 2020)

It is disconcerting to see that the same issues presented 16 years ago (D'Amico et al., 2005) are still the prominent ones today, even though industry leaders and experts (Cisco, 2014; Northcutt, 2014; Shackleford, 2012) warned us that work-load would only increase as more technologies (e.g. phones, virtual environments, wearables, etc.) connect to the Internet (i.e. The Internet of Things) creating more entryways for cyber criminals. Further they warned, that over time security network analysts would find it even more difficult to detect events and defend their environment.

It is clear that not much has been done over the years to resolve or to even improve the work challenges in the SOC, as some issues have only got worse, as seen in Figure, 1.2. Not surprising, 60 percent of the professionals surveyed in 2020 say the stress related burn-out of working in the SOC has caused them to consider changing and/or leaving their current job/position (Ponemon, 2020).

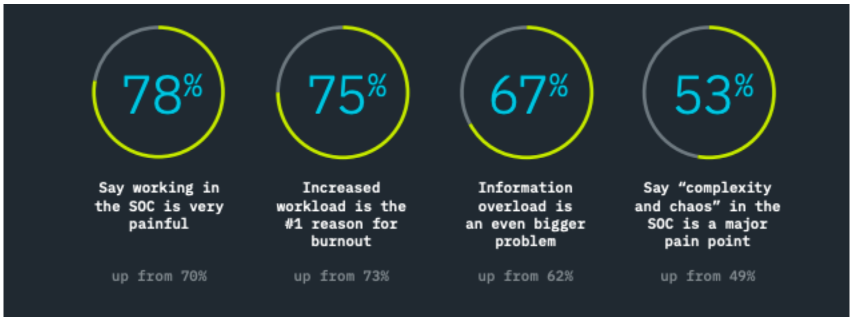


Figure 1.2. The percent of change between 2019 and 2020 per SOC issue (Ponemon, 2020)

In efforts to increase overall cyber performance effectiveness in the SOC, enterprises are urged and/or required to follow best practices for computer security (Newhouse et al., 2017; Petersen et al., 2020). That includes being able to identify individuals with the specific knowledge, skills, and abilities (KSA) needed to support defense of the enterprise (Huq, 2015; Newhouse et al., 2017). The majority of the human factors research (i.e., human/analyst centered) in cyber defense, aiming to identify the KSA requirements has investigated cyber performance differences between groups (e.g., teams) of individuals. Most of that research is experimental, where researchers investigated how cognition affected cyber performance when manipulating the network environment (Champion, Rajivan, Cooke, & Jariwala, 2012; Sawyer et al., 2015). There is very little research seeking to measure individual differences (i.e., between-subject variance) in basic human cognition (i.e., cognitive abilities) and how much they contribute to cyber performance success under naturally occurring circumstances (Cronbach, 1957). As a result, the cognitive abilities that mostly contribute to cyber performance success is still unknown. Many of the current issues seen in the SOC (Ponemon, 2020) can be better addressed if cognitive ability is identified.

1.3 Scope

The scope of the project was to identify the individual differences in cognitive ability that contributes to cyber performance for the role of Incident Responder. The Incident Responder work-role has been scientifically studied which offers enough evidence for pinpointing the cognitive ability that underlines performance in that particular role only. A primary task of the Incident Responder is to operate an incident detection system (IDS) (D'Amico et al., 2005; Goodall, Lutters, & Komlodi, 2009b), therefore the human-computer interactions in a simulated IDS (Funke et al., 2016) was collected from 16 cyber analysts from the Purdue University campus over the summer of 2020. Only the cyber analysts that had expertise using an IDS were included in the study. To quantify cognitive ability, the cyber analysts performed three direct cognitive measures that cover basic visual attention ability only. The measures included spatial attention (i.e., orienting, vigilance, executive function), working memory capacity, and executive functions (i.e., task-switching, inhibition and updating). Personality, emotions, mood and auditory research are out of scope, since they were seldom posited by the human-factors in cyber defense research to contribute to cyber performance. Lastly, the current study is an individual differences study (Cronbach, 1957) that quantitatively measures performance between-subjects rather than between groups. Therefore, the study uses correlational methods to accomplish that goal. The data collected from the IDS and from the cognitive ability measures is all that was required to detect the cognitive ability contributions to cyber performance success.

1.4 Research Question

The research question for the current study was: How do individual differences in cognitive ability contribute to the cyber performance of network security analysts? That question involved investigating which predictor variables have the strongest correlation with cyber performance, and if they can further predict cyber performance. The following is the hypothesis:

H_1: Individual differences in attention control predict the cyber performance of network security analysts.

1.5 Significance

The current study significantly contributes to the human factors research in the cyber security domain. That area of research aims to improve cyber defending through discovering more efficient strategies for defending networks against cyber attacks. It seeks improved training methods for analysts. It also designs network defense tools to aid human cognitive ability (Champion et al., 2012; Goodall, 2011; Vieane et al., 2016), and of most importance to the current study it investigates specific human factors that contribute to cyber performance. Even with those research efforts, the cognitive ability contribution to cyber performance has not yet been identified. The research proposed here aims to fill that gap in the human factors research (Dawson & Thomson, 2018) by scientifically identifying the individual differences in cognitive ability that contribute to cyber defense success. The research is important for several reasons.

The first reason for why the individual differences study is significant is because it transfers cognitive ability measures (i.e., attention control) developed in a laboratory, to the performance in a real-world cyber defense task. The successful transfer increases the predictive and ecological validity of those measures. It is of great importance because a main goal of cognitive sciences is to bridge the gap between the laboratory and real life (Holleman, Hooge, Kemner, & Hessels, 2020). Many studies that do attempt the transfer, do not use ecologically valid criterion representing the complexity of a real-world task (Holleman et al., 2020).

The next reason the study is important is for recruitment purposes. The cyber-security workforce (Ponemon, 2020) has expressed the challenge of recruiting and retaining individuals with the right knowledge, skills and abilities (Newhouse et al., 2017). One big reason for that challenge is that the cognitive abilities that directly contribute specifically to cyber knowledge and to skill (Motowildo et al., 1997; Schmidt, Hunter, & Outerbridge, 1986) are still unknown, and more or less assumed. As a result, when cyber-recruiters seek out a screening method to aid in their search for more cognitively equipped analysts, they either come up empty handed or with inadequate measures of cognitive ability, that have not yet been scientifically identified or associated with cyber performance (S. G. Campbell, O'Rourke, & Bunting, 2015; Trippe, Moriarty, Russell, Carretta, & Beatty, 2014). Therefore, many employers are currently using selection methods that are not very effective, and/or do not predict how the individual will perform (Schmidt & Hunter, 2004). The findings from the study will therefore directly benefit recruiters and employers in their search for more effective cyber analysts, as they will be able to screen for the scientifically identified cognitive abilities that are of most importance to cyber defense success. For example, individuals with a high cognitive ability can be selected for which to build a more efficient cyber security workforce, but only if the cognitive ability is associated with and/or predicts cyber performance. The research proposed here offers stable and reliable cognitive ability measures (i.e., predictors) that can predict cyber performance for hiring procedures, which can considerably increase the efficiency of cyber defending over time (Schmidt & Hunter, 1998).

One last reason for why the current research is important is that the findings could aid in the redesigning of incident detection systems and displays so that cyber analysts at any level of cognitive ability can still perform well. Many of the current software-tool designs do not take into consideration the limits of human cognition (Champion et al., 2012; Goodall et al., 2009b; R. S. Gutzwiller et al., 2016; Wickens, 2008), resulting in poor cyber defense performance and burn-out (Champion et al., 2012; Ponemon, 2020).

1.6 Assumptions

The assumptions for this study include:

- The cyber task is relevant and typical of a network analysts' job function.
- The cognitive tests used in this study are reliable and valid tests.
- The participants completed the measures presented to them to the best of their ability.
- The participants that performed the cyber task were indeed cyber analysts.
- The cognitive tasks chosen for this study were appropriate to answer the research questions.

1.7 Limitations

The limitations for this study include:

- The study was limited by the participants cooperation to perform the cognitive tests, and cyber tasks to completion and/or to correctness.
- The study was limited to 16 volunteer network analysts available who are affiliated with Purdue University, West Lafayette campus, over the summer of 2019.
- The study was limited by the reliability and validity of the measures.
- The study was limited by the time allotted for graduate research.
- The study was limited by the facility in which the research was conducted. As it was conducted in the back of a cyber laboratory, during the summer of 2019.
- The study was limited by funds available to pay participants for their time and effort.

1.8 Delimitations

The delimitations for this study include:

- The cyber task was limited to a small number of alerts (40) that a network analyst might encounter.
- The study was limited to just one task environment. Cyber analysts commonly use a variety of software to aid in their investigations.
- The study was limited to just one display that the analyst would use during cyber defense. While cyber analysts frequently shift between multiple displays.
- The study measured only three cognitive abilities while there are many other cognitive variables that contribute to the differences in cyber performance scores.

1.9 Summary

The chapter provided the scope, significance, research question, assumptions, limitations, delimitations, definitions, and other background information for the research project.

CHAPTER 2. REVIEW OF LITERATURE

In order to identify the cognitive ability contribution to cyber performance, first the important tasks associated with a specific cyber defense work-role must be identified, along with the human attributes that are likely to be important for success in that role. Next, the behaviors that differentiate the successful from the less successful cyber performers must be identified. Lastly, the human factors that affect cyber performance must be discovered. That process is said to allow for pinpointing/predicting the cognitive abilities that likely contribute most to the identified behaviors of cyber performance (Motowildo et al., 1997; Sauce & Matzel, 2013).

Accordingly, the chapter begins with an overview of the empirically identified work-role requirements specifically for the Incident Responder, along with the associated cognitive work-tasks and mental challenges in which they face. That is followed with an overview of the human factors research in cyber defense specific to investigating the cognitive factors that contribute to cyber performance in incident detection system-like tasks. The chapter addresses that through the more popular cognitive theory/frameworks that have been applied to study cyber performance, such as Situational Awareness (CyberSA) (Endsley, 1995), Signal Detection Theory (D. M. Green, Swets, et al., 1966) and the Hybrid Space Framework (Jøsok et al., 2019). Each section gives an overview of the specific framework followed by the identified human factors posited to influence cyber performance. Additionally, the identified behavioral differences that contribute to more successful cyber performance is highlighted. The reasoning for the literature review to address the cognitive human-factors and behaviors that effect and/or contribute to cyber performance in an IDS, is that the main goal of individual differences research is to make sense of the effects that have already been observed (Cronbach, 1957; Sauce & Matzel, 2013). The chapter will show that the correlational research is limited in regards to the identification of individual differences in cognitive abilities, that can influence cyber performance.

2.1 Individual Differences in Job Performance

Cronbach (1957) famously made the distinction between the experimental research stream verses the correlational research stream. He wrote that the experimental approach studies the variance among treatments. It is also used for determining underlying causes for behavior (Sauce & Matzel, 2013). While the correlational approach is used to investigate the variance among individuals (Cronbach, 1957). The correlational approach is also used to study which way variables interact in a population to produce differences in a behavior as, "causes of variation of a behavior" (Sauce & Matzel, 2013). The current study applies the correlational research approach and its methods under the well-established individual differences framework (Sauce & Matzel, 2013). The individual differences framework recognizes that individuals differ in behavior such as in personality traits, genetics, and cognitive ability. The individual differences in cognitive ability and how it influences job performance behavior is of importance to the current study. The association between the two is one of the most established associations in the literature (Hunter, 1983; Motowildo et al., 1997; Schmidt et al., 1986). The frameworks of job performance (Hunter, 1983; Motowildo et al., 1997; Schmidt et al., 1986) all hold the common idea that individual differences in cognitive ability directly impacts job performance (e.g., a work sample), and indirectly impacts job performance through job knowledge, as seen in Figure, 2.2. Those interactions displayed in the causal model become important when determining the cognitive abilities that contribute to cyber performance.

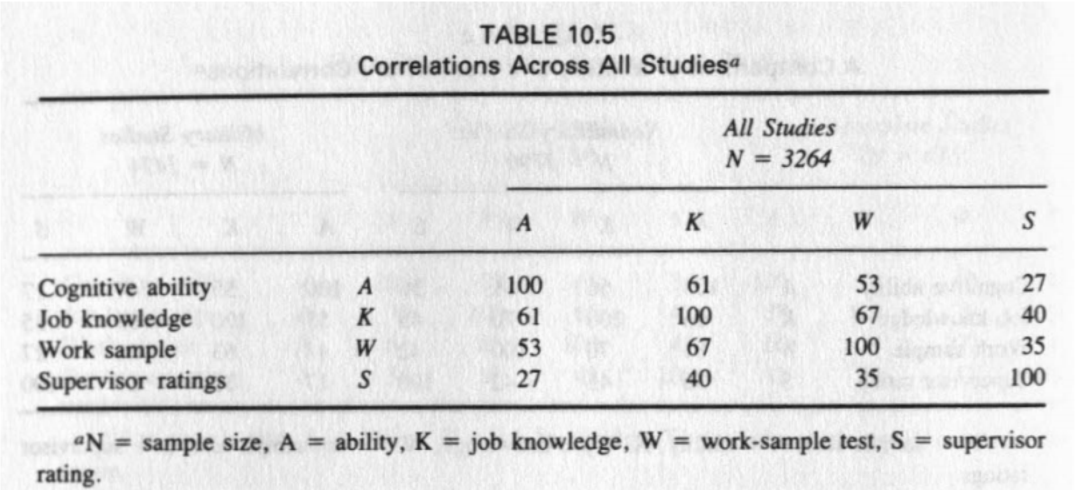


TABLE 10.5
Correlations Across All Studies[a]

		All Studies N = 3264			
		A	*K*	*W*	*S*
Cognitive ability	*A*	100	61	53	27
Job knowledge	*K*	61	100	67	40
Work sample	*W*	53	67	100	35
Supervisor ratings	*S*	27	40	35	100

[a]N = sample size, A = ability, K = job knowledge, W = work-sample test, S = supervisor rating.

Figure 2.1. Correlations Across all Studies

Hunter (1983) put the framework to use and investigated the relationships among the three variables in various occupations using a large meta-analysis (n = 3264 cases, n = 14 studies). The researcher only included previous studies with similar cognitive ability predictors and similar

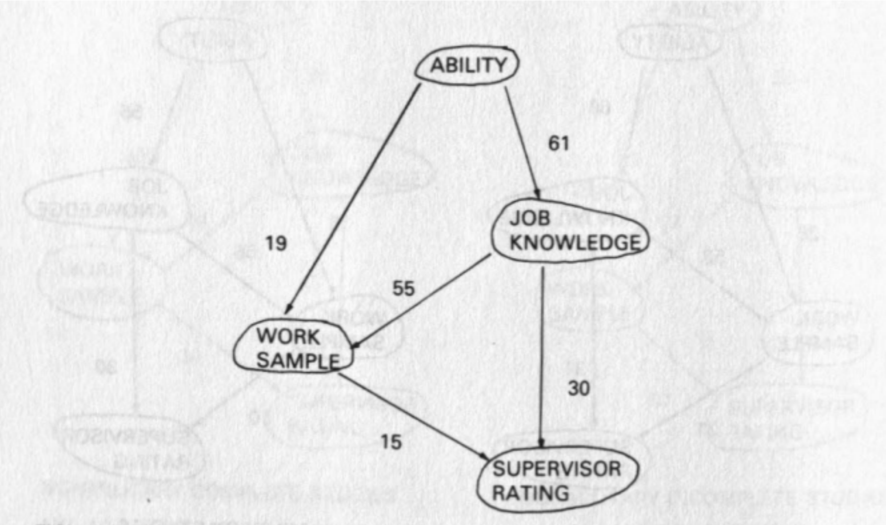


FIG. 10.3. The causal model that fits the average correlations across all studies, $N = 3264$.

Figure 2.2. A Causal Model for Cognitive Ability, Knowledge and Performance (Hunter, 1983)

criterion measures (e.g., direct verses subjective job performance measures). The researcher found a similar average reliability between the three variables across the studies. His results are seen Figure, 2.1, and his causal model can be seen in Figure, 2.2. The high correlation of (r = 0.53) between cognitive ability and performance (i.e., work-samples) from the reliability estimates in Figure, 2.1 comes from a direct causal impact (r = 0.19) and also from an indirect effect of (r = 0.61 − 0.55 = 0.34) through job knowledge, as seen in the causal model in Figure, 2.2. Those correlations show that cognitive ability directly contributes to work performance itself and to knowledge.

In a more present meta-analysis consisting of over 32,000 employees in 515 widely diverse jobs, Schmidt and Hunter (2004) found that general mental ability (GMA) is the highest predictor of job performance. The researchers reported that the validity of the GMA for predicting job performance was (r = 0.58) for professional-managerial jobs, (r = 0.56) for high-level complex technical jobs, and further declines as complexity/skill of the job lessens. The researchers conclude that the GMA is a better predictor of job performance over other predictors because of its higher association with knowledge. Individuals with higher cognitive ability obtain knowledge quicker and perform better. When applied to the current study, cognitive ability should be related to cyber performance to the extent that the job-task calls for or depends on the specific abilities (Hunter, 1983). When considering the empirical research (Schmidt & Hunter, 2004, 1998; Schmidt et al., 1986) and theoretical support (Motowildo et al., 1997) for the causal model (Hunter, 1983) and the large, repeatedly discovered association (r = 0.53) between cognitive ability and performance, it is surprising how little it has been studied in the cyber performance domain for job selection purposes when both knowledge and cognitive ability contribute to cyber performance and to each other.

There has been recent gravitation towards the use of cognitive ability measures for cyber recruitment purposes (Svenmarck, 2020). That is because the demand for cyber security personnel for the various cyber defense work-roles (Petersen et al., 2020) is currently higher than the amount of students that graduate in computer science (Svenmarck, 2020). One solution has been the utilization of cognitive ability tasks beyond the usual education requirement of a computer science degree in order to increase the recruitment base (Svenmarck, 2020). However, only a handful of cognitive ability tests have been devised for recruitment of cyber personnel, such as the Armed Services Vocational Aptitude Battery (ASVAB) (Trippe et al., 2014), or the Cyber Aptitude and Talent Assessment (CATA) (S. G. Campbell et al., 2015). Those popular assessments apply cognitive ability measures that have not yet been scientifically identified in association with cyber performance from a specific cyber work-task and/or within a specific work-role, which eliminates its ability to predict job performance. In order to identify the specific cognitive abilities that most affect cyber performance, first the critical work functions/processes, tasks, and cognitive challenges of a specific cyber defense work-role must be empirically identified. That will allow for a definition of job-specific cyber performance (Motowildo et al., 1997) that will lead us to a representative work-sample for which cognitive ability contributions can be properly assessed. The following section identifies the important knowledge, skills and tasks for the Incident Responder work-role (Petersen et al., 2020).

2.1.1 Knowledge, Skills and Tasks for Cyber Defense

In 2017, the National Initiative for the Cyber-security Workforce Framework (NICE) was created by the National Institute of Standards and Technology (NIST) in the U.S. Department of Commerce (Newhouse et al., 2017). The framework has been recently revised (Petersen et al., 2020). The NICE framework serves as the most popular go-to guide for how to identify, recruit, develop, and retain cyber-security talent for a more efficient cyber workforce. The former NICE framework (Newhouse et al., 2017) identified the tasks, knowledge, skills, and abilities for strengthening the cyber-security of an organization (i.e., KSAs). Those factors are continuously updated (Petersen et al., 2020) by coordinating with government, academic, and industry partners. For example, the revised version of the NICE framework seen in Figure, 2.3 omits the ability factor.

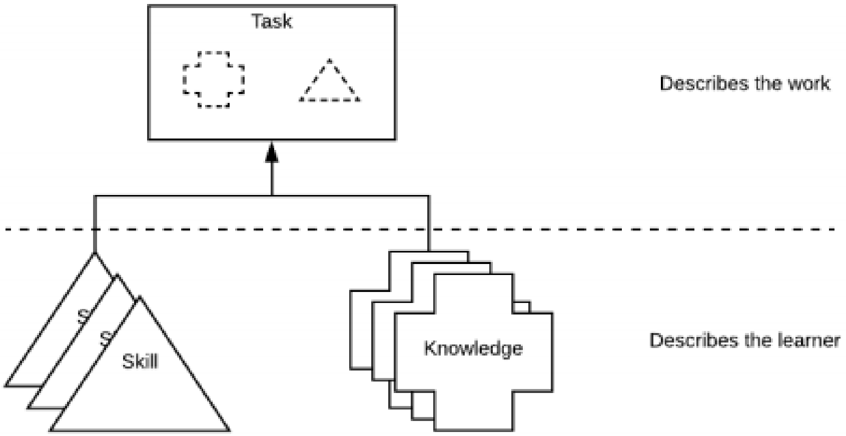


Figure 2.3. The NICE Framework: TKS (Petersen et al., 2020)

The NICE framework (Petersen et al., 2020) now identifies the skills, knowledge requirements and work-tasks (TSKs) for all of the various cyber defense work-roles. The skills from the framework that are needed for cyber security are defined as, “The capacity to perform an observable action (i.e., describes what the learner can do) (Petersen et al., 2020, p.5). The tasks from the framework are described as, “An activity that is directed toward the achievement of organizational objective (i.e., describes the work to be done) (Petersen et al., 2020, p.4). Lastly, knowledge is described as, “ A retrievable set of concepts within memory” (Petersen et al., 2020, p.5). The framework provides no information regarding the required cognitive abilities that directly affect knowledge, skill, and contribute to more successful cyber-defending (Hunter, 1983; Motowildo et al., 1997; Schmidt & Hunter, 2004). For example, the required skills for Cyber Defense Incident Responder are listed from the NICE Framework as:

- Skill of identifying, capturing, containing, and reporting malware.
- Skill in preserving evidence integrity according to standard operating procedures or national standards.
- Skill in securing network communications.
- Skill in recognizing and categorizing types of vulnerabilities and associated attacks.
- Skill in protecting a network against malware. (e.g., NIPS, anti-malware, restrict/prevent external devices, spam filters).
- Skill in performing damage assessments.
- Skill in using security event correlation tools.
- Skill to design incident response for cloud service models.

The two required abilities for the Cyber Defense Incident Responder role that have been removed from the most recent NICE Framework (Newhouse et al., 2017) are:

- Ability to design incident response for cloud service models.
- Ability to apply techniques for detecting host and network-based intrusions using intrusion detection technologies.

The human factors research in cyber defense (Andrade et al., 2018; Jones, Namin, & Armstrong, 2018; Kokkonen & Puuska, 2018) has been attempting to fill in the gap of the NICE framework in regards to what the skill requirements and previously described ability (Newhouse et al., 2017) further entail. For instance, one might ask, what exactly are the desired human attributes and cognitive abilities that likely contribute to more success in the cyber defense Incident Responder role? In efforts to begin answering that question, cognitive task analysis (CTA) for cyber defense work performance (D'Amico et al., 2005; D'Amico & Whitley, 2008; R. S. Gutzwiller et al., 2016) has proven helpful. An understanding of the primary tasks (i.e., behavior) and mental challenges that analysts face during incident detection work allows us to start thinking about associated cognitive abilities (i.e., predictors) that might aid in that work environment (Motowildo et al., 1997; Sauce & Matzel, 2013). The next section discusses the work-role of the Incident Responder, followed by the associated cognitive challenges from CTA research.

2.1.2 The Role of the Cyber Analyst: Incident Responder

The role of Incident Responder is vaguely described in the NICE framework as,“investigates, analyzes, and responds to cyber incidents within the network environment or enclave” (Newhouse et al., 2017, p.111). Fortunately, for purposes of the current study, the cognitive requirements/processes and tasks that are essential to the Incident Responder role have been generally identified in cognitive task analysis (CTA) research (Champion et al., 2012; D’Amico et al., 2005; Goodall et al., 2009b; R. S. Gutzwiller et al., 2016; Mahoney et al., 2010; Zhong, 2016) and the Incident Responder work-role has been further described in the following research.

First, D’Amico et al. (2005) conducted one of the most widely cited CTA studies in cyber defense to present, that looked at the entire work process of information security. Their CTA used 41 network security analysts from commercial and government organizations. The analysis of the data revealed a generalized workflow for the various roles of cyber analyst, to include three stages:

- Event detection: monitoring and detection
- Situation assessment: analysis
- Threat assessment: response

The event detection stage (stage 1) for the triage analysis role (i.e., Incident Responder) is most relevant to the current study, additionally shown in Figure, 2.4. In the ‘detection stage’ (i.e., triage analysis) the main role of the cyber analyst is described as being able to quickly detect and escalate suspicious events (i.e., true-positive threats) that are located within large arrays of sensor data (e.g., IDS alerts, firewall logs, OS audit trails, vulnerability reports and packet dumps) for further investigation. The data within the incident detection system; a common network software that automatically alerts the analyst to signs of network intrusions, is known to contain mostly

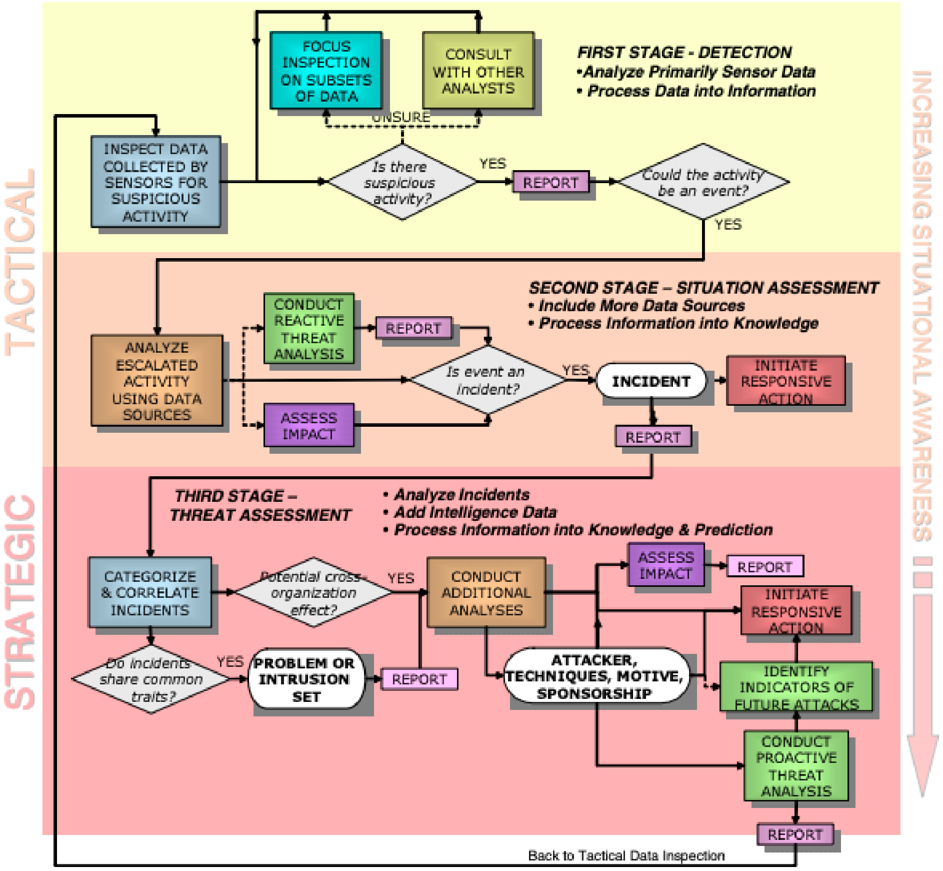

TACTICAL
STRATEGIC
INCREASING SITUATIONAL AWARENESS
FIRST STAGE - DETECTION
•Analyze Primarily Sensor Data
• Process Data into Information
FOCUS INSPECTION ON SUBSETS OF DATA
CONSULT WITH OTHER ANALYSTS
UNSURE
INSPECT DATA COLLECTED BY SENSORS FOR SUSPICIOUS ACTIVITY
Is there suspicious activity?
YES
REPORT
Could the activity be an event?
YES
SECOND STAGE – SITUATION ASSESSMENT
• Include More Data Sources
• Process Information into Knowledge
ANALYZE ESCALATED ACTIVITY USING DATA SOURCES
CONDUCT REACTIVE THREAT ANALYSIS
REPORT
Is event an incident?
YES
INCIDENT
INITIATE RESPONSIVE ACTION
ASSESS IMPACT
REPORT
THIRD STAGE – THREAT ASSESSMENT
• Analyze Incidents
• Add Intelligence Data
• Process Information into Knowledge & Prediction
CATEGORIZE & CORRELATE INCIDENTS
Potential cross-organization effect?
YES
CONDUCT ADDITIONAL ANALYSES
ASSESS IMPACT
REPORT
INITIATE RESPONSIVE ACTION
Do incidents share common traits?
YES
PROBLEM OR INTRUSION SET
REPORT
ATTACKER, TECHNIQUES, MOTIVE, SPONSORSHIP
IDENTIFY INDICATORS OF FUTURE ATTACKS
CONDUCT PROACTIVE THREAT ANALYSIS
REPORT
Back to Tactical Data Inspection

Figure 2.4. Cyber Defense Work-flow (D'Amico et al., 2005)

false-positive alerts (Champion et al., 2012; Goodall et al., 2009b). The analysts are described as, 'weeding through it' in search of actual threats (D'Amico et al., 2005). More specific to reactive incident detection tasks and activities, Goodall, Lutters, and Komlodi (2009a); Goodall et al. (2009b) presented four common stages for the role of Network Intrusion Detection Analyst as follows:

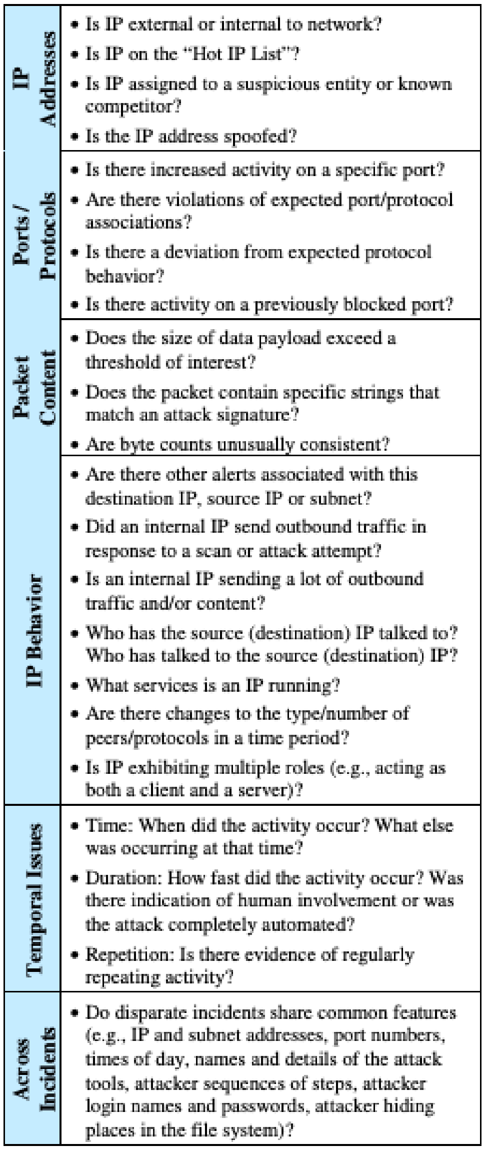
- Monitoring the network for events: looking for indications of anomalous or malicious activity
- Triaging an event: prioritizing an alert and/or threat, which can lead to a temporary stopgap measure
- Analysis of an event: the analyst uses various tools to determine if the alert is a false-alarm or not

- Response to an event: a stopgap measure, such as patching a system

All analysts participating in their study followed the four stage process when performing incident detection work. Accordingly, analysts monitor the network for threats through an incident detection system (IDS). When alerted to a potential event the analyst will triage it, analyse it through network data tools (e.g., Tripwire, Snort) and then respond to it. The researchers described the work as dynamic, event-driven, and reactive (e.g., to an IDS alert or event). Additional CTA research (D'Amico et al., 2005; Paul & Whitley, 2013) has identified the associated analytic questions that Incident Responders must ask themselves as they move through the process towards making the threat decisions, as seen in Figure, 2.5.

The questions are said to be complex (e.g., multi-step), requiring significant cyber-domain knowledge and situated expertise (Goodall et al., 2009a) in order to successfully answer them given the information and tools provided in the network environment. The researchers state that, "Incident detection systems (IDS) cue human analysts and provide them with targeted (but often overwhelming) network data to aid in their investigation. Because of this human-centered nature of incident detection work, it is especially important to understand the types of expertise needed and the ways this expertise is acquired. Understanding these will help incident detection tool designers, trainers, managers, and incident detection practitioners" (Goodall et al., 2009b, p.11). Although investigating the knowledge and expertise contribution to cyber performance is important, Schmidt and Hunter (2004) found that job expertise and education only have a small correlation of ($r = 0.18$ and $r = 0.10$) with job performance. Identifying the cognitive ability contribution to cyber performance will aid in understanding more about how expertise is acquired. For instance, if the analyst will be able to achieve it in a timely manner.

To date, the analytical questions and general work-role findings of the Incident Responder (D'Amico et al., 2005; Goodall et al., 2009a) have been widely cited and upheld in the human factors research. Further CTA research (Goodall et al., 2009a; R. S. Gutzwiller et al., 2016; Paul & Whitley, 2013; Zhong, 2016) reveal more fine-grained cognitive processes of triage analysis and additional work-roles and tasks specifically for the Incident Responder. That allows us a view into the Incident Responders behaviors and decision-making process, and to then further question the cognitive ability contribution that might aid in a more successful performance in that role (Sauce & Matzel, 2013).



IP Addresses	• Is IP external or internal to network? • Is IP on the "Hot IP List"? • Is IP assigned to a suspicious entity or known competitor? • Is the IP address spoofed?
Ports / Protocols	• Is there increased activity on a specific port? • Are there violations of expected port/protocol associations? • Is there a deviation from expected protocol behavior? • Is there activity on a previously blocked port?
Packet Content	• Does the size of data payload exceed a threshold of interest? • Does the packet contain specific strings that match an attack signature? • Are byte counts unusually consistent?
IP Behavior	• Are there other alerts associated with this destination IP, source IP or subnet? • Did an internal IP send outbound traffic in response to a scan or attack attempt? • Is an internal IP sending a lot of outbound traffic and/or content? • Who has the source (destination) IP talked to? Who has talked to the source (destination) IP? • What services is an IP running? • Are there changes to the type/number of peers/protocols in a time period? • Is IP exhibiting multiple roles (e.g., acting as both a client and a server)?
Temporal Issues	• Time: When did the activity occur? What else was occurring at that time? • Duration: How fast did the activity occur? Was there indication of human involvement or was the attack completely automated? • Repetition: Is there evidence of regularly repeating activity?
Across Incidents	• Do disparate incidents share common features (e.g., IP and subnet addresses, port numbers, times of day, names and details of the attack tools, attacker sequences of steps, attacker login names and passwords, attacker hiding places in the file system)?

Figure 2.5. Analysis Questions (D'Amico et al., 2005)

2.1.3 Cognitive Challenges and Primary Tasks of the Cyber Analyst

The cognitive task analysis (CTA) research identifying the skills, knowledge, and abilities of the cyber analyst have found a similar compilation of cognitive challenges that analysts encounter during incident detection work, that have mostly not improved over the last 16 years (Ponemon, 2020). One persisting cognitive challenge for the Incident Responder is to manage the increasing information-overload issue in the security operations center (SOC) (Cisco, 2020;

D'Amico et al., 2005; Ponemon, 2020) without allowing it to affect primary task performance. The primary cognitive task of the Incident Responder and commonly noted challenge due to information overload (Cisco, 2020; Ponemon, 2020) includes the formation of mental models per incident/event (D'Amico et al., 2005; Goodall et al., 2009a; R. S. Gutzwiller et al., 2016; Wickens, 2008). The primary task (i.e., an information search and evidence collection task) entails searching through high volumes of various network data for evidence associated with each incident. That requires shifting through the use of many different vendor software/tools (Cisco, 2020; Goodall, 2011), although the amount of vendor tools used in the SOC has reduced since 2017. Now, 86 percent of organisations use between one and 20 vendors (Cisco, 2020). The data of importance discovered from each tool must then be mentally fused or correlated together (i.e., data-fusion) into a mental model without using external aids. The events are typically correlated by timestamps, IP addresses, host names, and port numbers, on an ad hoc basis (i.e., reconstructing the events timeline) to decide if malicious activity occurred. The additional research (Champion et al., 2012; Goodall et al., 2009b; R. S. Gutzwiller et al., 2016) studying cognitive tasks of the Incident Responder, and current industry research (Cisco, 2020; Ponemon, 2020) all report the significant data-overload challenge in association with the primary task, where analysts literally face thousands of alert investigations per day. R. S. Gutzwiller et al. (2016) described the analysts from their CTA as "taxed," as they work on incidents from "cradle to the grave". That type of working environment leads to burnout and to the inability to retain cyber personnel (Ponemon, 2020), although higher cognitive ability may lessen those effects (Sauce & Matzel, 2013).

Another highlighted cognitive challenge to the Incident Responder role is that analysts must make threat decisions in an IDS often under uncertainty or ambiguity. That is said to be due to limitations with network tools in their ability to present information that the analyst needs to make proper decisions (Champion et al., 2012; D'Amico et al., 2005; Goodall et al., 2009b; Ponemon, 2020; Tyworth, Giacobe, & Mancuso, 2012). The specific tool limitations posited in that research are that they provide incomplete sensor data, and do not allow for visibility into the attack surface. That causes the analyst to switch between the use of many different tools. When the analyst does not have all of the evidence that they need to make the threat decision, visual attention research (Wolfe, Horowitz, & Kenner, 2005) suggests that it can cause extensive searches for the missing evidence (i.e., target information) (Wickens, Gutzwiller, & Santamaria, 2015; Wolfe et al., 2005). The extent of the search behavior likely depends on the cognitive ability of the individual (Wolfe et al., 2005) and/or their perceived task difficulty, work-habits, effort, etc. (Motowildo et al., 1997). Unfortunately, the specific individual differences in cognitive ability that mostly contribute to overcoming the cognitive challenges in IDS work is again not considered in the cognitive task analysis research.

2.1.4 Human Attention Limits in Cyber Defense

The cognitive task analysis (CTA) research (D'Amico et al., 2005; Goodall et al., 2009b) and current industry research (Ponemon, 2020), identifies that human capacity limits are being exceeded in cyber defense work. In light of that fact there is a sprouting body of visualization research (Tyworth et al., 2012; Vieane et al., 2017, 2016) that aims to reduce cognitive load (i.e., data-overload) in order to improve analyst performance. To give just two examples; through designing visual aids in IDS software that can guide attention (Posner & DiGirolamo, 1998) to areas of importance (R. S. Gutzwiller et al., 2016), and through the development of automation software. However, the aids and software tools are said to not always enhance cyber performance since they seldom take into consideration the analysts goals, information needs (Goodall et al., 2009b; R. S. Gutzwiller et al., 2016; Wickens, 2008), and the limits of human cognition (Champion et al., 2012). For example, incident detection systems (i.e., automation software) still produce high amounts of false-alarms (Ponemon, 2020) which cognitively overloads the analyst and hinders cyber performance (Champion et al., 2012; Ponemon, 2020).

Further, experimental research (Mancuso, Minotra, Giacobe, McNeese, & Tyworth, 2012) demonstrates how difficult it is to actually measure the effects of a cyber interface/tool on cognition and because of that difficulty, the researchers point out that software designers advertising that a cyber-tool improved cognition is probably not supported by research. If the cognitive ability that contributes most to cyber performance (e.g. IDS performance) has not yet been identified, one must ask how it is possible to design effective aids and network tools to ease cognitive processing in the first place. Although, from the research presented here (Vieane et al., 2016), the task displays that have been designed with consideration of human cognition limitations still appear to increase the cyber performance success.

There are individual differences (i.e., variance) in human capacity limits that determines the effect that a visual aid has on cyber performance, but interactions between the individuals cognitive ability and the new aid/tool (i.e., treatment variable) is often neglected in cyber visualization research (Vieane et al., 2017, 2016). The individual differences in cognition can be measured through traditional cognitive ability measures in association with the cyber performance scores, with the ultimate goal being to increase cyber performance (i.e., payoff) for those that scored lower in cognitive ability. That idea can be seen in Figure 2.6 from Cronbach (1957).

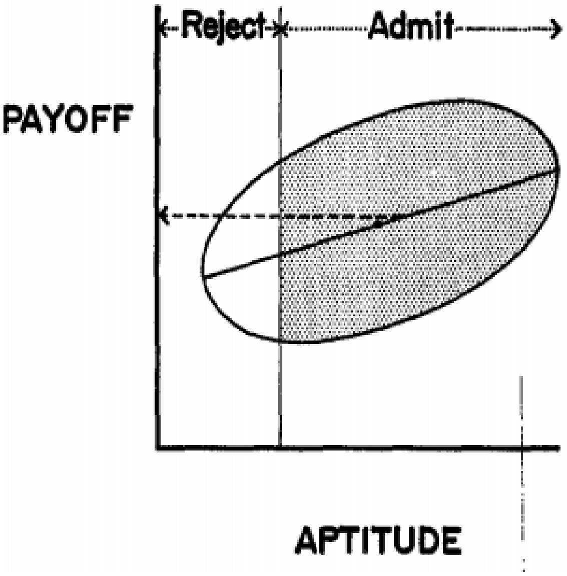


Figure 2.6. The Figure Shows that the Admission of Students with High Scores on a Relevant Aptitude are Admitted, which Raises the Payoff for the Institution (Cronbach, 1957)

Cronbach (1957) said that the correlationist applies a fixed treatment and looks for aptitudes that maximize the slope of the payoff (i.e., cyber performance). In cyber defense selection, the organization will take the analysts with high scores on the relevant aptitude (e.g., cognitive ability) and thus raise the payoff for the organization. First the cognitive abilities that contribute most to differences in cyber performance must be identified (Cronbach, 1957).

2.1.5 Cognitive Task Analysis Summary

The cognitive task analysis (CTA) research scientifically identified the core work processes, cognitive challenges, and higher-order cognitive requirements most fully for the Incident Responder. That lays the ground-work towards identifying the actual cognitive abilities that drive successful performance in the Incident Responder work-role. Individual differences in cognitive ability can mitigate the primary cognitive challenges, discussed from CTA research: the formation of multiple mental-models and data-fusion (De Jong, 2000). However, we still do not know which cognitive ability contributes most to its success. Studying individual differences in cognitive abilities, and identifying which ones associate most with differences in cyber performance is one way to find solutions to mitigate human capacity effects (Champion et al., 2012) on cyber performance (Sauce & Matzel, 2013).

The repeated finding from the CTA research that security needs are not being met because incident detection work is not compatible with basic human-information processing abilities (R. S. Gutzwiller et al., 2016; Wickens, 2008) further highlights the need for identifying the cognitive ability that most contributes to cyber performance. If we know what the cognitive abilities are that contribute most to successful cyber performance in specific work-tasks then more effective technologies/software can be devised/redesigned so that individual differences no longer show meaningful variation in performance (Guastello, Shircel, Malon, & Timm, 2015). The individual differences framework (Cronbach, 1957; Motowildo et al., 1997) must be applied in cyber defense/human factors research in order to identify them. The next section discusses the synthetic task environments that are widely used to study human factors in cyber defense.

2.2 Synthetic Task Environments

Synthetic task environments (STE), are primarily used by the human factors researchers in cyber defense for ultimately improving cyber performance. STEs are often used for studying aspects of human cognition (Champion et al., 2012; Cronbach, 1957; Greenlee et al., 2016), team performance/collaboration, cyber defense training, and for improving software aids and network tools (Vieane et al., 2017, 2016). The STEs in cyber defense research range from simulating a real-world network environment such as an intrusion detection system (IDS) (Ben-Asher & Gonzalez, 2015; Champion et al., 2012; Funke et al., 2016; Mancuso et al., 2012), to simulating very specific incident response tasks, such as a network monitoring task (McIntire et al., 2013; Sawyer et al., 2015). One reason for the wide use of a STE in human factors research is because cyber defense environments are inaccessible (Mancuso et al., 2012). Another reason for its use is because it offers a controlled environment with less confounds from which a more reliable measure of cyber performance can be obtained.

Unfortunately, many of the human performance studies in cyber defense either use a STE that is not similar to any real-world cyber task, and/or a new one is introduced that has not been validated in further research. That reduces how representative it is to how cyber analysts behave. The behaviors (i.e., human-computer interactions) of the analyst are usually recorded as they operate in the STE environment, and then the human-computer interactions of interest to the researcher are analyzed depending on the applied human-performance framework or model.

Therefore, the next three sections discuss the main theories and frameworks that have been applied to study cognitive processes of cyber analysts while they defended networks in a STE, beginning with Situational Awareness (i.e., CyberSA) (Endsley, 1995), followed by Signal Detection Theory (D. M. Green et al., 1966) and ending with the Hybrid Space Framework (Jøsok et al., 2019). Each section includes the associated human factors and human computer interactions posited to contribute to cyber defense performance. Of most importance to the current study, the patterns of behaviors that differentiate the effective from the less effective cyber performers in an incident detection environment are identified. That allowed for pinpointing the cognitive ability variables that may account for cyber performance (Motowildo et al., 1997; Sauce & Matzel, 2013).

2.3 CyberSA Framework

One human performance model (i.e., pre-decision making model) that has been extensively utilized in the human factors research in cyber defense is the situational awareness (SA) model by Endsley (1995). The SA model is shown in Figure, 2.7.

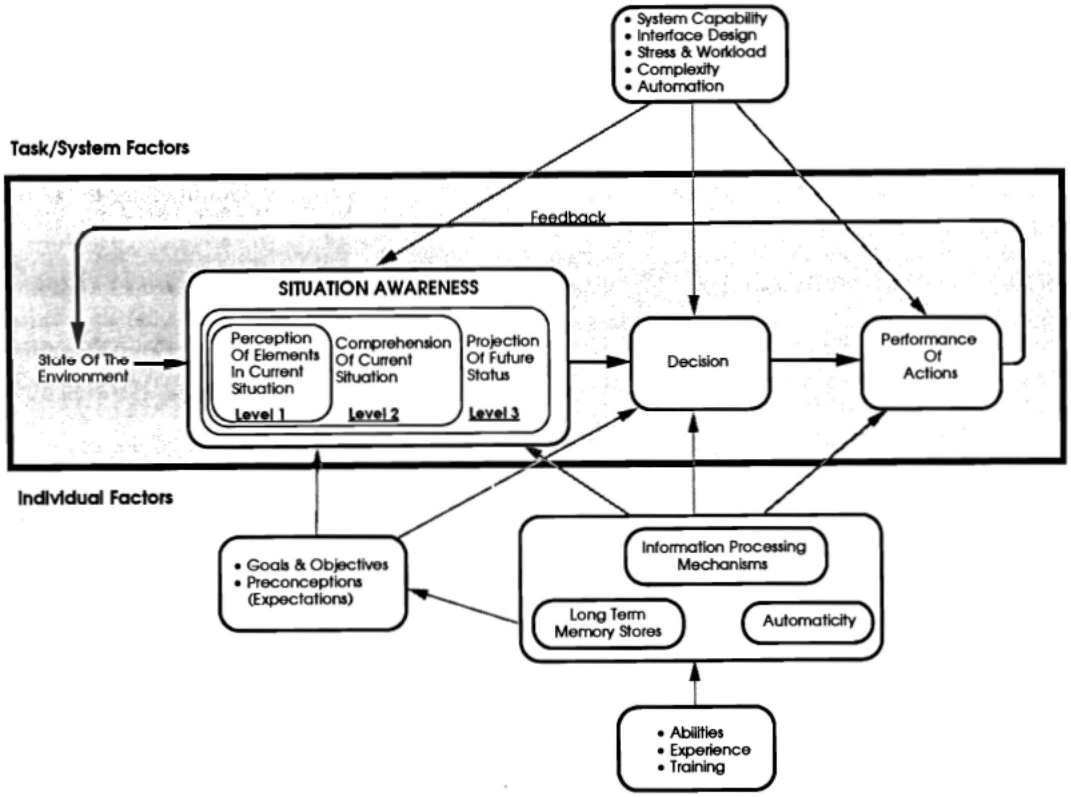


Figure 2.7. Situational Awareness Model (Endsley, 1995) Showing the Attention and Working Memory Contribution to SA

Situational awareness is formally defined as, “The perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future” (Endsley, 1995, p.12). In other words, SA is viewed as achieving a ‘state of knowledge’ or an understanding of what is going on in a dynamic environment before a decision can be reached. The three cognitive processes of SA from its definition include:

- Perception: perception of critical elements of information in the environment (Level 1)
- Comprehension: acquiring the meaning of the information in the environment by combining the elements (Level 2)
- Projection: a projection of actions based on the perception and comprehension that was gained in the environment (Level 3)

The information-processing mechanisms and innate abilities of the individual, such as attention and working memory as shown in Figure, 2.7 have been generally recognized to contribute to achieving SA (R. Gutzwiller, 2007). When using the model as a conceptual tool to obtain a measure (i.e., a level) of SA, subjective (i.e., indirect) measures are often applied, such as self-rating questionnaires (e.g., SART), observer-rating scores (Champion et al., 2012), or confidence scales (Evangelopoulou & Johnson, 2015). Those subjective measures of cognition have draw-backs since they measure how the individual perceived their own SA rather than their actual (i.e., observed) SA, and often they do not correlate with actual performance scores (Champion et al., 2012; Evangelopoulou & Johnson, 2015; Mancuso et al., 2012). In addition, probing methods are applied such as the situation awareness global assessment technique (SAGAT) (Endsley, 1995; Mancuso et al., 2012), where the individual is asked comprehension questions and is given stop-and-go feedback about what is really happening in the environment (Zhong, 2016). The stop-and-go measures and probing of any type in real-time for assessing SA can be disruptive to the on-going task and can interfere with correlations between cognitive ability (e.g., SA) and performance (Mancuso et al., 2012).

When the SA model is properly applied to cyber defense (CyberSA), which entails actually assessing the levels of SA, the SA achievement level is said to depend on both the capability of the tool/system to present the required information as well as the cognitive ability of the individual (R. S. Gutzwiller et al., 2016). The CTA research (Champion et al., 2012; D'Amico et al., 2005; R. S. Gutzwiller et al., 2016) has recognized the issues on both sides of that relationship. For example, the limitations with the network tools/software, the high data volume issue, and the associated data-overload experienced by cyber analysts, but the actual cognitive ability contributions have not been fully addressed.

The following section of CyberSA research (Champion et al., 2012; Evangelopoulou & Johnson, 2015; Paul & Whitley, 2013; Stevens-Adams et al., 2013) reports the consequences for the failing relationship between the analyst and the tools/software, where it finds CyberSA is difficult to maintain. More specifically, the projection level has been found difficult to achieve within a cyber environment (Evangelopoulou & Johnson, 2015; R. S. Gutzwiller et al., 2016) with most cyber performance studies (Ben-Asher & Gonzalez, 2015; Champion et al., 2012; Dutt, Ahn, & Gonzalez, 2013; Stevens-Adams et al., 2013) finding that either CyberSA and/or task performance is low. The next subsections explore the proposed human factors that have been found to affect cyber performance and/or CyberSA, and the methods that were used to identify the factors are discussed.

2.3.1 Cognitive Workload

A human factor of interest in the cyber defense research is the cognitive-workload factor. Cognitive workload can be defined as, " The level of attentional resources required to meet both objective and subjective criteria, which may be mediated by task demands, external support, and past experience" (Young & Stanton, 2001, p.10). The theory behind cognitive work-load is that humans have a limited capacity and/or limited resources to process information (Kahneman, Treisman, et al., 1984). That means that as task difficulty increases so does mental workload, and once the environmental demands exceed the individuals capacity, task performance then declines. A high work-load is often associated with negative outcomes such as fatigue, frustration and stress, and it can affect cyber performance. A high cognitive work-load has been reported in the cognitive task analysis studies and in the CyberSA research in cyber defense (Champion et al., 2012; R. S. Gutzwiller et al., 2016; Vieane et al., 2017) as a cognitive challenge for Incident Responders; partly stemming from data-overload. The cognitive-workload has been commonly measured through the NASA Task Load Index (TLX) questionnaire of perceived work-load, in direct association with cyber performance like in the following study.

Champion et al. (2012) conducted one of the first cyber defense experiments to show how a higher work-load negatively affects team CyberSA performance (Endsley, 1995). The researchers used eight teams of three analysts each (n = 23 Cadets) from a cyber security student group. The team performance was assessed from the percent of correctly identified attack paths, from classifications of individual attack alerts in an IDS, and through a self-assessment of CyberSA. Furthermore, the analysts completed two subjective measures; the NASA Task Load Index (TLX) questionnaire and a confidence measure. Those measures were collected twice from two different trials, with the second trial containing 47 additional events (false-positive alerts).

The researchers found that performance scores were all quite low with only 60 percent of the events correctly classified, and only 47 percent of the teams correctly identifying the attack paths. That finding lined up with CyberSA which was scored as moderate-to-low, in that the teams reported being only "somewhat aware" of what was occurring in the task. However, the finding of poor performance from measures all around did not line up with the teams high self-reported confidence in their categorization of alerts (70 percent confident). Additionally, a somewhat high cognitive load was reported from the NASA Task Load Index; in the high six range for both trials. There was only one significant difference ($p = 0.05$) in performance between the two trials with more incorrectly classified events (i.e., reconnaissance events only) seen in the trial in which the the false-positives were added. The performance decline was 16 percent from trial one to trial two. The researchers were actually only able to show that higher false-alarms caused a decline in performance (i.e., cognitive overload) in association with CyberSA when the amount of false-positive events in the IDS was increased. There were no other discovered differences between trials for any other measures including the NASA-TLX measure of subjective work-load.

The study highlights the effects of high cognitive demand (i.e., high data volume) of the IDS environment, a mental challenge posited by cognitive task analysis research (D'Amico et al., 2005; R. S. Gutzwiller et al., 2016) and it shows the negative effect that it has on CyberSA. However, it does not show the relative importance of the "cognitive capacity variable" to making successful threat decisions. For instance, if the researchers had investigated the scores between individuals rather than solely between teams, we would learn if their subjective cognitive work-load (NASA-TLX) mitigated the effect (Sauce & Matzel, 2013), but capacity measures that correlate directly with performance would be required.

Further, the study highlights the inaccuracies of self-assessed performance (Van Zandt, 2000) in the cyber defense domain, which can be seen in the analysts' low accuracy performance in relation to their self-reported over-confidence per alert decision. The discrepancy is repeatedly observed in cyber defense studies (Ben-Asher & Gonzalez, 2015) and could be related to the ambiguous nature of alert classifications (D'Amico et al., 2005; Tyworth et al., 2012). The study also shows the limitations of using 'proportions correct' only to assess IDS performance. Here we miss out in knowing if the difference in performance between the two different trials differ in sensitivity, response bias, or both. We do know that with the addition of false-positive alarms the analyst may have altered their criterion (i.e., response bias) for classifying threats

(Lerman et al., 2010), but like most studies under CyberSA, the decision strategy (i.e., response bias) was not considered.

Newer research from Wright-Patterson Air Force Base (Greenlee et al., 2016), further measured the high-cognitive demand in an IDS environment because its effects (i.e., stress and burnout) may cause an analyst to lose their situational awareness, but unlike the previous study (Champion et al., 2012) specific network analysis tasks (i.e., face validated) related to cyber work-roles were used and compared. The researchers investigated the cognitive work-load and stress from operating in an IDS for two different network analysis tasks; triage ($n = 27$) and escalation ($n = 46$). The researchers applied two subjective measures; the Dundee Stress State Questionnaire (DSSQ) used to measure task engagement, distress, and worry, while the NASA-TLX was used to measure cognitive work-load. In both tasks the mental demand and effort was rated as significantly greater than 50 ($p < 0.001$), meaning it was high. The Multiple Resources Questionnaire (MRQ), a work-load subscale further showed that triage analysis demanded significantly more short-term memory, spatial attentive processing, spatial emergent processing, and visual lexical processing, compared with other subscales ($p < 0.05$) and in comparison to escalation analysis. From the findings of the study the researchers point to sustained spatial attention and/or vigilance as possibly playing a major role in triage analysis. Although the exploratory study used college-aged individuals without any cyber experience, the findings serve as clues to the cognitive ability underlying more successful cyber performance in an IDS (Sauce & Matzel, 2013).

2.3.2 Working Memory

Working memory (WM) is a core cognitive factor that underlies CyberSA performance (Endsley, 1995), and may contribute to a more successful performance in cyber defending. For example, working memory is involved in the formation of mental models and data-fusion; the primary cognitive challenges identified in CTA research for IDS work (D'Amico et al., 2005; R. S. Gutzwiller et al., 2016; Wickens, 2008). Working memory was first conceptualized by Baddeley and Hitch (1974), and discussed as a limited capacity system or resource that can manipulate and update stored (i.e., maintained) information over a short amount of time (i.e., short term memory). It is considered a core cognitive process that is used in every day tasks that are not automatically performed (Unsworth & Engle, 2007). For example, it is used for actively maintaining information such as phone numbers or current task goals. Additionally, the stored information can be manipulated with the deployment of executive processes (i.e., attention control processes). For example, repeating a series of digits backwards requires both storage and manipulation. The tasks designed to measure working memory are said to measure both attention control process (e.g., inhibition and updating) and its storage capacity (Baddeley & Hitch, 1974; Engle & Kane, 2004; Miyake & Friedman, 2012; Shipstead, Harrison, & Engle, 2015). Working memory is often considered in cognitive models of cyber behavior (Dutt et al., 2013).

Cyber defense researchers (Ben-Asher, Oltramari, Erbacher, & Gonzalez, 2015; Cranford et al., 2020; Dutt et al., 2013; Yuan, Li, Rusconi, & Aljaffan, 2017) have been using cognitive modeling (i.e., cognitive decision models) to simulate human behaviors when operating in a computing environment and to better understand the involved human cognitive processes/abilities such as perception, memory, and attention (Endsley, 1995) in efforts to improve software interfaces and CyberSA performance. For example, (Dutt et al., 2013) used working memory as a core component in their computational cognitive model of human behavior. The model is based on the Instance-Based Learning Theory (IBL) to expected cyber defender behavior in various attack and defend situations. The IBL factors included the level of prior knowledge (i.e., experiences stored in a simulated analysts working memory) and the analysts risk-tolerance level (i.e., the point at which the analyst decides to classify a sequence of events as a cyber-attack or wait for more convincing evidence). CyberSA performance was measured by accuracy (i.e., precision and recall) and timeliness of the analysts decision to stop the attack. One finding from the study was that the model predicted that exposure to more threat experiences, increased the models working memory ability to detect threats. However, other causes can be even more important as causes of variation in cyber performance that must be identified through the individual differences framework (Sauce & Matzel, 2013) using behavioral data from actual analysts.

More recent research has used eye tracking (Yuan et al., 2017) to better understand cognitive processes of individuals when interacting with security systems for contributing to human performance models, but again the actual cognitive ability behind the behavior is not identified. The cognitive modeling approach for studying cyber performance was pointed out as being a good alternative to drawing conclusions from samples of undergraduate students like the experimental studies (Ben-Asher & Gonzalez, 2015; Champion et al., 2012) have done due to the difficulty of obtaining actual cyber analysts for research purposes.

2.3.3 Task-Switching

Situational Awareness is considered a main precursor to decision-making, but it can degrade with fatigue and stress, and it can be negatively affected by task interruptions (i.e., task-switching) and distractions (Endsley, 1995), as seen in the cyber defense environment (Vieane et al., 2017, 2016). Task-switching refers to the time that it takes to shift attention (i.e., ease of transitioning) to different tasks (e.g., rules and task-sets) (Miyake & Friedman, 2012; Miyake et al., 2000). The theory around set-shifting tasks is that the cognitive control processes (i.e. attention control) are significantly more involved when the individual switches between tasks (ABAB) in an incongruent trial, in comparison to a congruent trial where the individual repeats the same task (AAAA) (Jersild, 1927; Monsell, 2003). The two trials are used to isolate the executive-control processes (a incongruent trial) from non-executive-control processes (a congruent trial). The performance on the task is measured in response time (D. A. Allport, Styles, & Hsieh, 1994; Jersild, 1927; Rogers & Monsell, 1995; Vandierendonck, Liefooghe, & Verbruggen, 2010) and error. The performance difference between those two trials represents the cognitive control processes engaging during switching, which is reflected in a shift-cost. The findings from a more recent study (Kortschot et al., 2018) that investigated attentional switching in the cyber defense domain suggests that task switching may affect cyber performance, however real-world tasks were not used and/or the participants did not have prior experience and knowledge of the experimental platform that was applied; as similarly seen in the following end-user study.

Vieane et al. (2016) investigated whether coordinated displays (i.e., displays that automatically link relevant event information by timestamp across databases), in comparison to standard displays improved an analysts IDS performance. The researchers conducted a true experiment (N = 46) and used a t-test to measure the between group differences in time to completion. The novice participants with no cyber experience in the uncoordinated condition took significantly longer (M = 86.21 minutes, SE = 4.62 minutes) to finish the alerts than did participants in the coordinated condition (M = 43.61 minutes, SE = 3.08 minutes), t (42) = 7.67, p < .001, Cohen's d = 2.31. They discovered that coordinated displays nearly doubled CyberSA performance efficiency (i.e, time) in detecting network threats with the simulated environment. The finding suggests that task switching may affect cyber performance.

Further, Vieane et al. (2017) found that task interruptions such as answering e-mails during the classification of attacks, significantly caused a performance decline for 13 novice with no cyber experience. The participants ranged in age from 18 to 33. The participants correctly classified most of the alerts, however the alerts that were interrupted took significantly (p = .005) longer (M = 72.02 s, SE = 5.57 s) for participants to complete in comparison to alerts that were not interrupted (M = 63.44 s, SE = 4.85 s). The previous two studies (Vieane et al., 2017, 2016) therefore suggest that task-switching negatively affected cyber performance.

2.3.4 CyberSA Summary

Cognitive work-load and/or capacity theory (Champion et al., 2012) and task-switching (Vieane et al., 2017, 2016) was identified as affecting CyberSA, while the other human factors proposed to influence CyberSA are spatial attention and vigilance (Greenlee et al., 2016), and higher working memory in association with more experience and knowledge (Dutt et al., 2013). However, the CyberSA research fails to take into account the individual and their unique set of knowledge, experiences, and cognitive ability that influences CyberSA. For instance, the basic cognitive ability contribution that is considered essential to making decisions under the CyberSA conceptual model such as attention and working memory (Endsley, 1995; R. Gutzwiller, 2007) is yet to be empirically validated in association with cyber performance under an individual differences framework. One reason for that is because most of the human factors research addresses the cyber performance of analysts' in teams, while the individual differences (i.e., causes in variation) contribution to performance is not considered.

Another reason may be because there is limited ways in which to measure CyberSA directly without interfering with on-going cyber performance (Mancuso et al., 2012). In order to appropriately study the differences between analysts that contribute to cyber performance and/or to CyberSA, it is recommended that researchers use measures that are direct and reliable (e.g., applied in more than one study) and further, that measures are high in sensitivity and low in measurement error. However, the subjective measures applied under CyberSA framework are not of the essence (Salmon, Stanton, Walker, & Green, 2006). One last limitation to using subjective CyberSA measures are that they are known to reflect only the behavior of the participants and not their 'internal processing of information' (Endsley, 1995).

Overall, Champion et al. (2012) put it best when they questioned how SA transfers to a cyber environment. They said, "How does this apply with the cyber world in which perception is limited to what a computer can convey through the monitor, where space is seemingly infinite, and comprehension is shared between the computer and analyst?" (Champion et al., 2012, p.1). From that description of CyberSA, the Signal Detection Theory (D. M. Green et al., 1966; N. A. Macmillan & Creelman, 1990) discussed next could be a better choice for measuring cyber performance, especially when many of the choices that analysts make are ambiguous. SDT is another popular human performance model that takes into account the analysts internal processing and/or neural noise (Lynn & Barrett, 2014) and it also considers the external noise stemming from the task (e.g., how different the alerts are from each other) in overall performance.

2.4 Signal Detection Theory Framework

To better assess cyber performance under uncertainty, the Signal Detection Theory (SDT) (D. M. Green et al., 1966; N. Macmillan & Creelman, 2005) has been applied in the human factors research (Ben-Asher & Gonzalez, 2015). Originally, SDT was developed to assess the binary responses (i.e., classification behavior) from discrimination tasks that cause perceptual uncertainty (Fechner, Howes, & Boring, 1966). Fechner et al. (1966) was first to propose SDT, when he explained how to scale mental experiences (i.e., psychophysical scaling). For that he presented how to scale/measure subjective sensations of heaviness caused from two different weights by asking the participant which one is heavier. He further explained that individuals will require a certain amount of heaviness that will correspond to their neural noise before they can detect a change in the weight. In other words, the binary decisions ('yes' or 'no') that humans make, are said to require a certain internal threshold of evidence (i.e., neural noise) or confidence before they can be made. Individuals will differ in how much evidence they require. His same logic applies to recognition memory decisions (Kantner & Lindsay, 2012; N. A. Macmillan & Creelman, 1990; Mickes, Wais, & Wixted, 2009), eye-witness identifications (Wixted, Mickes, Dunn, Clark, & Wells, 2016), weather forecasting (Harvey Jr, Hammond, Lusk, & Mross, 1992), and in cyber defense decision making (Ben-Asher & Gonzalez, 2015). The two main components to decision making under SDT are:

- Sensitivity/discriminability
- Response Bias

Applied to cyber defense for clarity, the SDT asks when presented with two alerts (e.g., a false-positive threat and a true-positive threat) over a series of trials, can an analyst detect the sensation of a threat (i.e., the signal) from the sensation of a false-alarm (i.e., the noise). That is known as the individuals discrimination (i.e., sensitivity) ability. The SDT also asks if the individual is more inclined to respond to alerts as false-positives or as hits when faced with the ambiguous/subjective information presented by cyber tools. That is known as the response bias (i.e., decision strategy). Even though cognitive factors (i.e., internal factors) such as neural noise and changes in attention may affect the response (Stanislaw & Todorov, 1999), the SDT does not identify any specific cognitive processes or abilities that may contribute to making a decision. SDT quantifies the overall sensitivity and response bias that are confounded in an accuracy score (Lynn & Barrett, 2014), for which to precisely study an individuals performance.

The following section will cover the cyber research using SDT metrics and/or the psycho-physical methods from SDT, as measures of cyber performance. First, the section will cover the cyber vigilance research using SDT methods, followed by the research using SDT to measure performance differences based on knowledge. The research investigating the response bias of cyber analysts is also addressed in the section. Of most importance to the current study, the following research further identifies important cognitive variables contributing to cyber performance.

2.4.1 Spatial Attention

Attention is a human factor that underlies cyber performance and the achievement of CyberSA (Endsley, 1995). It is difficult to define attention because of its many different types, interpretations, and associations with other cognitive processes, but most definitions refer to it as the cognitive mechanism(s) (i.e., a gateway) that allow certain information (i.e., objects, features and locations) (Desimone & Duncan, 1995) to be more fully processed than the information not selected (Cohen, Cavanagh, Chun, & Nakayama, 2012). Many cognitive theories concur that human behavior is controlled by top-down attention (i.e., voluntary attention or goal-directed attention) and bottom-up attention (i.e., involuntary attention, reflexive, or stimulus-driven attention) (Desimone & Duncan, 1995; Itti & Koch, 2001; Koch & Tsuchiya, 2007). Attention that is guided by top-down processes is using instructions, prior knowledge and/or goals stored in working memory to control behavior (Desimone & Duncan, 1995) in order to accomplish a task. Attention that is guided by bottom-up processes is automatically captured by objects or features (Theeuwes & Burger, 1998) because of their distinctiveness or uniqueness (i.e., salience, color, location) (Theeuwes & Burger, 1998), or because the information is relevant to the observer's intentions or goals (Folk, Remington, & Johnston, 1992). Those two attention types interact in ways that either benefit or hinder the individual during tasks. For instance, attention can automatically or involuntarily shift (i.e., orient) with bottom-up attention to a flickering light even if it is irrelevant to ones current task-goals. In that potentially hindering situation, research (Posner & DiGirolamo, 1998) suggests that bottom-up attention is competing or conflicting with top-down attentional goals, as consistently reflected in longer RT; observed in conflict tasks (Eriksen, 1995; Yantis & Johnston, 1990). The resulting deployment of inhibition (i.e., attention control) on the distracting information by top-down attention (i.e., executive functions) is also important for allowing the individual to stay on task (Desimone & Duncan, 1995; Engle & Kane, 2004; Folk et al., 1992; Sawaki & Luck, 2010; Yantis, 1992), especially when the primary task load is high (Lavie, 1995). The amount interference experienced from irrelevant information depends on individual differences in attention control (Engle & Kane, 2004; Miyake & Friedman, 2012), and/or their capacity under perceptual load theory (Lavie, 1995).

Researchers (Emmanuel, McClain, Matzen, & Forsythe, 2015) from Sandia National Laboratory have investigated attention differences in novice from expert analysts in a conference contribution. They applied the visual search paradigm to cyber security since the tasks and the tools used by analysts require them to visually scan through large amounts of text-centric data (e.g., logs), in order to search for events of interest. The eye-tracking activity (i.e., attention allocation) of 12 novice and expert analysts was measured while they completed a static log analysis task in a variety of commonly used cyber tools (e.g., Network Miner, Wireshark, ENCASE Enterprise, AccessData Registry, PDF Dissector, and Hex Workshop). The researchers found that the novice users took three times longer to find the target of interest in the log and were often distracted by 'extraneous text' that caught their attention. Although the study is a conference contribution and incomplete, it still serves as another clue for the current study towards identifying the basic cognitive ability, such as visual attention (i.e., visual search) that contributes most to cyber performance differences.

2.4.2 Vigilance

The following section discusses how vigilance, a lower-level cognitive ability affects cyber performance through the use of Signal Detection methods. First, visual attention is fully described as follows. Posner (1990), proposed a widely accepted model of visual attention with three brain networks/constructs:

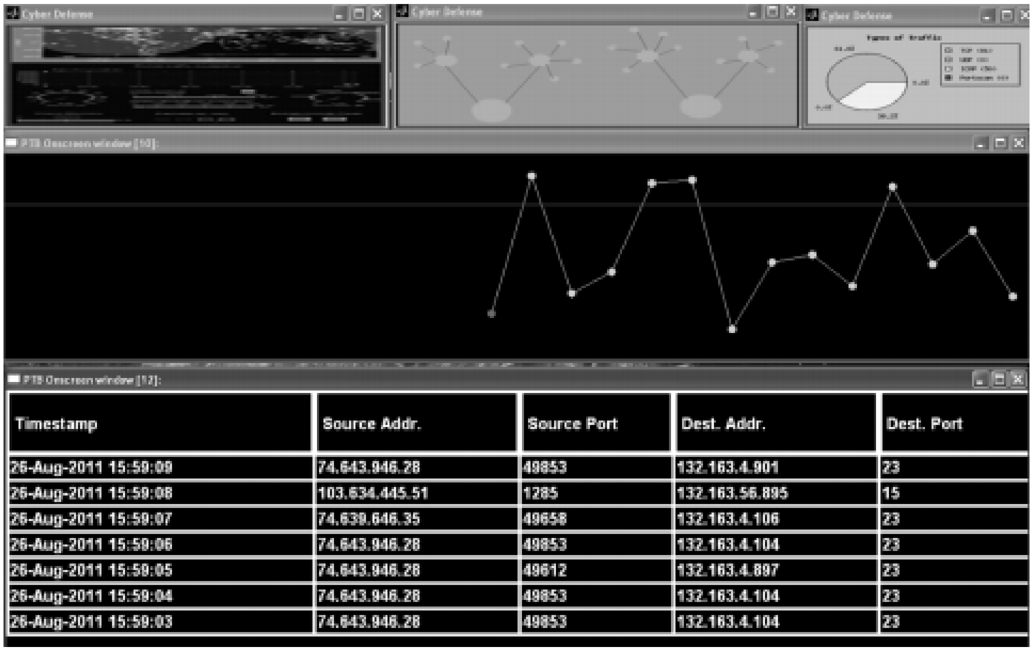
- Alerting
- Orienting
- Executive Control

According to the model, the orienting network (Desimone & Duncan, 1995; Itti & Koch, 2001), allows the individual to move/orient their "spotlight of attention" to areas of interest in the environment. The alerting network (i.e., vigilance) includes the individuals ability to stay in an alerted state or to maintain their readiness to respond during a task. Lastly, the executive control network (Engle & Kane, 2004; Miyake & Friedman, 2012) (i.e., attention control network) includes the inhibition processes that are used to resolve conflict such as competing attentional task demands. The conflict for example, could be distracting information (e.g., internal or external conflict) occurring during the task. Further, each network is associated with distinct anatomy in the brain, but they stay functionally connected (Fan et al., 2005; Fan, McCandliss, Sommer, Raz, & Posner, 2002; McConnell & Shore, 2011). The small correlations observed between the networks still allow for them to be studied separately (i.e., diverse or domain specific) (Fan et al., 2002).

Cyber defense researchers in human factors (McIntire et al., 2013; Sawyer et al., 2015), have questioned if a cyber defense monitoring task- an aspect of IDS work (Goodall et al., 2009b) is associated with vigilance tasks. That is because a sustained attention or vigilance in tasks, over a long duration (i.e., time-on-task) is associated with a vigilance decline (i.e., a vigilance decrement) (Hitchcock et al., 2003). A vigilance decline or decrement (i.e., attention failure) in cyber defense results in individuals missing critical events/threats that often times lead to very negative consequences and/or costs (McIntire et al., 2013; Sawyer et al., 2015). The cyber defense research (McIntire et al., 2013; Sawyer et al., 2015) manipulated and then investigated psycho-metric factors under SDT, such as event rate (i.e., the amount of events to be monitored) and signal probability (i.e., probability of the critical signal/target occurring) in cyber monitoring tasks and then assessed if the performance associated with vigilance.

(McIntire et al., 2013) was the first to investigate vigilance through eye-tracking in a controlled cyber task, under methods of SDT. The analysts (n = 20) were given a response task that included monitoring a graph display for unusual levels of network activity. The task also required monitoring of a text display for suspicious Internet protocol (IP) addresses and port combinations coming into the network. Their task dashboard can be seen in Figure 2.8.

The analyst pushed a key-board button when they recognized the suspicious IP address that was memorized prior to performing the cyber task. In both displays the critical signal event rate (i.e. suspicious IP) was a low five percent. It is difficult to achieve a representative event rate in cyber defense research since real-world true-positive threat encounters are very low; somewhere around one to five percent depending on the network environment (Funke et al., 2016; Goodall et al., 2009b). The performance efficiency was then assessed from the percentage of correct signals detected (i.e., percent hits), calculated every 10 minutes (i.e., 10 minute epochs)



Timestamp	Source Addr.	Source Port	Dest. Addr.	Dest. Port
26-Aug-2011 15:59:09	74.643.946.28	49853	132.163.4.901	23
26-Aug-2011 15:59:08	103.634.445.51	1285	132.163.56.895	15
26-Aug-2011 15:59:07	74.639.646.35	49658	132.163.4.106	23
26-Aug-2011 15:59:06	74.643.946.28	49853	132.163.4.104	23
26-Aug-2011 15:59:05	74.643.946.28	49612	132.163.4.897	23
26-Aug-2011 15:59:04	74.643.946.28	49853	132.163.4.104	23
26-Aug-2011 15:59:03	74.643.946.28	49853	132.163.4.104	23

Figure 2.8. The Cyber Monitoring Task Dashboard (McIntire et al., 2013)

over the 40 minute trials because the critical signals were set to appear at those times. The researchers found that the measured oculometrics (i.e., eye tracking) to assess operator vigilance, significantly correlated with changes in cyber performance over time on task (p = .05). The negative correlations between the percent hits and whether there was a decrease in performance (i.e., decrement group) or not (i.e., no decrement group) can be seen in Figure 2.9.

As attention (i.e., vigilance) decreased (e.g., slower eye blinking) cyber performance (i.e., hit-rate) declined; over time. That was said to represent the vigilance decrement (Hitchcock et al., 2003). The researchers suggest shortening the time spent in a cyber task in order to reduce the decrement. Individual differences in vigilance can be further investigated in association with cyber performance, through traditional vigilance tasks instead of eye tracking.

Years later Sawyer et al. (2015) conducted an experimental study (n = 24) to determine if cyber tasks are associated with traditional vigilance tasks. They used a cyber response task (i.e., discrimination task) that was similar to the previous study (McIntire et al., 2013). The critical signal for detection was a simple visual discrimination between the IP address and associated communication port- usually flagged by an IDS system (Goodall et al., 2009b). The critical

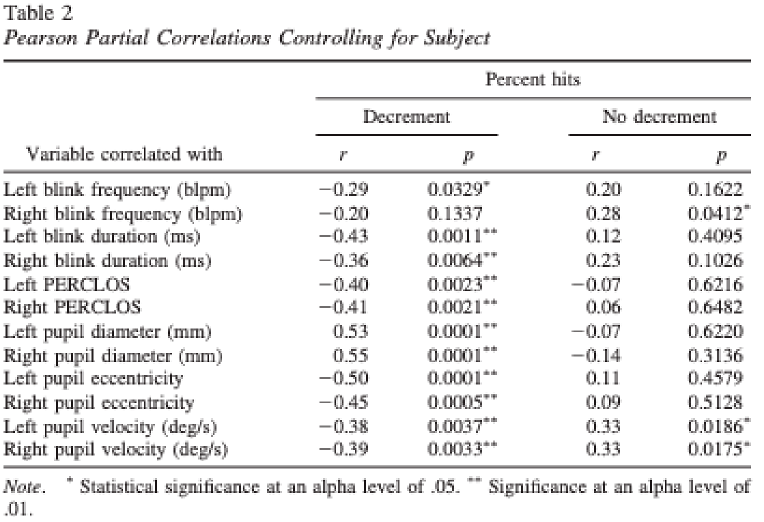


Table 2
Pearson Partial Correlations Controlling for Subject

	Percent hits			
	Decrement		No decrement	
Variable correlated with	*r*	*p*	*r*	*p*
Left blink frequency (blpm)	−0.29	0.0329*	0.20	0.1622
Right blink frequency (blpm)	−0.20	0.1337	0.28	0.0412*
Left blink duration (ms)	−0.43	0.0011**	0.12	0.4095
Right blink duration (ms)	−0.36	0.0064**	0.23	0.1026
Left PERCLOS	−0.40	0.0023**	−0.07	0.6216
Right PERCLOS	−0.41	0.0021**	0.06	0.6482
Left pupil diameter (mm)	0.53	0.0001**	−0.07	0.6220
Right pupil diameter (mm)	0.55	0.0001**	−0.14	0.3136
Left pupil eccentricity	−0.50	0.0001**	0.11	0.4579
Right pupil eccentricity	−0.45	0.0005**	0.09	0.5128
Left pupil velocity (deg/s)	−0.38	0.0037**	0.33	0.0186*
Right pupil velocity (deg/s)	−0.39	0.0033**	0.33	0.0175*

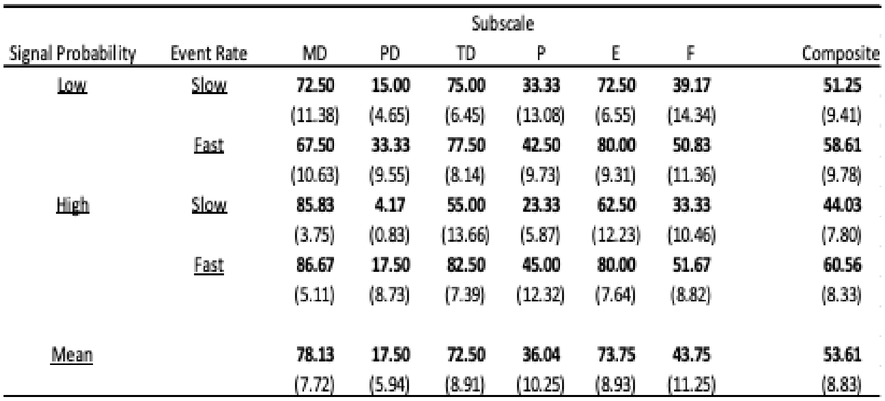
Note. * Statistical significance at an alpha level of .05. ** Significance at an alpha level of .01.

Figure 2.9. The Negative Correlations Between the Percent Hits and Whether There was a Decrease in Performance (i.e., Decrement Group) or not (i.e., no Decrement Group) (McIntire et al., 2013)

signal was again manipulated by high and low probability of its occurrence (i.e., signal probability), and the amount of events to be monitored (i.e., event rate) was also manipulated. After completing the cyber task, the participants mental demand (i.e., work-load) was measured through the NASA Task Load Index (NASA-TLX).

The researchers found that the detection rate for the critical signal was lower when it appeared less in the low signal probability (M = 72.14 percent) condition, when compared to the high signal probability condition (M = 83.14 percent) at (p = .05). Further, the mean detection scores (M = 88 percent) were higher in the slow event rate condition verses the fast event rate condition (M = 66 percent) at ($p < .001$). Then, the high self-reported mental workload (NASA-TLX) for the task was associated with a decline in performance in those conditions, as seen in Figure 2.10. The findings were said to suggest that the cyber task was quickly susceptible to the vigilance decrement.

However, one issue with the study was that the cyber task was manipulated into a 'vigilance task' to begin with even though performance research (Cronbach, 1957; Robertson & Kandola, 1982) warns scientists against that practice. Overall, the researchers were able to demonstrate that a simple cyber monitoring task (i.e., discrimination task) requires significant attentional demands under pressure, limited time and volume (Greenlee et al., 2016), and that the



		Subscale						
Signal Probability	Event Rate	MD	PD	TD	P	E	F	Composite
Low	Slow	**72.50**	**15.00**	**75.00**	**33.33**	**72.50**	**39.17**	**51.25**
		(11.38)	(4.65)	(6.45)	(13.08)	(6.55)	(14.34)	(9.41)
	Fast	**67.50**	**33.33**	**77.50**	**42.50**	**80.00**	**50.83**	**58.61**
		(10.63)	(9.55)	(8.14)	(9.73)	(9.31)	(11.36)	(9.78)
High	Slow	**85.83**	**4.17**	**55.00**	**23.33**	**62.50**	**33.33**	**44.03**
		(3.75)	(0.83)	(13.66)	(5.87)	(12.23)	(10.46)	(7.80)
	Fast	**86.67**	**17.50**	**82.50**	**45.00**	**80.00**	**51.67**	**60.56**
		(5.11)	(8.73)	(7.39)	(12.32)	(7.64)	(8.82)	(8.33)
Mean		**78.13**	**17.50**	**72.50**	**36.04**	**73.75**	**43.75**	**53.61**
		(7.72)	(5.94)	(8.91)	(10.25)	(8.93)	(11.25)	(8.83)

Note: Standard errors are in parentheses. Mean NASA Task Load Index (TLX) scores are listed for the subscales of Mental Demand (MD), Physical Demand (PD), Temporal Demand (TD), Performance (P), Effort (E), and Frustration (F).

Figure 2.10. Self-reported Mental Workload (NASA-TLX) for the Cyber Task (Sawyer et al., 2015)

higher work-load leads to missed true-positive threat detection. The task demands in their study are said to be comparable to the tasks in air-traffic control and medical monitoring, but cyber defense environments tend not to require monitoring beyond quick checks for suspicious activity (Goodall et al., 2009b). Furthermore, the researchers conclude that the analysts could not maintain a high performance because of, “diminished information processing resources, a situation that is arguably reflected in the high scores seen on the NASA TLX, especially in the Effort subscale” (Sawyer et al., 2015, p.161), but that would require a individual differences investigation.

In conclusion, the results from the experimental research (i.e., observing variation between groups) discussed here suggests that high vigilance is required to complete simple cyber tasks. Research is still needed to assess how individual differences in vigilance affects the cyber performance. That can be done through real-world, cyber defense tasks where analysts investigate rather than simply monitor (i.e., discriminate and respond) for threats. The monitoring task used in the previous two studies is usually performed by the IDS anyway (Goodall et al., 2009b). The following sections extends the SDT beyond discrimination tasks, to acknowledging that in IDS tasks analysts make their decisions through ambiguous choices. It also discusses the human-computer behaviors that have been discovered to differentiate a successful cyber group from the less successful through traditional Signal Detection Theory (D. M. Green et al., 1966; N. Macmillan & Creelman, 2005). Most significantly, the identified behavioral variables of more successful cyber defense can then be further predicted by cognitive abilities.

2.4.3 Knowledge

Knowledge is a factor that affects cognitive ability, skills and task performance (Motowildo et al., 1997). Cyber domain-knowledge and technical skills are well defined factors under the Cyber-security Workforce Framework (i.e., NICE) (Northcutt, 2016; Petersen et al., 2020) however, having individual knowledge or expertise alone does not guarantee success, as seen in the following study. (Ben-Asher & Gonzalez, 2015) first applied the Signal Detection Theory (SDT) to investigate how the domain-knowledge factor (e.g., expertise and practical knowledge) and cognitive ability contributes to cyber performance (i.e., information search and evidence collection) for the cyber expert in comparison to the novice (n = 55). The researchers designed a simplified IDS task that held multiple attack scenarios within it (e.g., a denial of service scenario, and a deface-website scenario, etc.). The IDS task produced an alert (network event) every 10 seconds, and the participants classified it as a threat or not. After the completion of each scenario they were also required to determine if the entire scenario was a cyber attack or not and to give the confidence in their decision. The researchers designed the IDS task so that novices (i.e., students from the general student population) could perform reasonably well, in which they noted vastly reduced its representation of a real-world cyber environment. A further limitation to their study was that the task experience was not the same for both groups, as the novice group participated in-person (60 minutes long) and was given 10 of the scenarios at random, while the expert took an online version of the task (25 minutes long), and was given only three of the 10 scenarios at random.

The researchers reported overall performance in regards to alert classifications in the IDS as low- near chance level, as seen in previous experimental studies (Champion et al., 2012) using a IDS simulator. The 20 cyber security professionals correctly detected 55 percent of the 20 malicious events with a 15 percent false detection rate (d' = 1.18), while the novice performed significantly worse (p = .024) at 45 percent (d' = .78). Under SDT a value of 0 means no detection ability, while a value of 3.00 and up reflects high sensitivity (N. Macmillan & Creelman, 2005). Therefore, the performance level indicates high task difficulty and/or a low internal sensitivity for detecting threats from non-threats for IDS work (N. A. Macmillan & Creelman, 1990). Further, the experts obtained a significantly higher hit-rate (i.e., correct classification of true-positives), and a lower false-alarm rate in the stealing confidential data scenario (t(240) = 4.105, p = < .001), and in the sniffer detected scenario (t(116) = 1.948, p = .054) to suggest knowledge influences cyber performance in only two out of four scenarios. The individual differences in cognitive ability (i.e., causes of variation in performance) that strongly affects knowledge (Motowildo et al., 1997), was not actually investigated in the study.

In a further analysis, in regards to the attack scenarios, the researchers cleverly applied SDT with a focus on utility (Lynn & Barrett, 2014) asking at what point do analysts perceive and then classify that an attack has occurred when given a sequence of network events leading up to an attack. They found that performance and confidence in the determination of the attack was the same between the novice and the expert, to which they suggest knowledge did not play a significant role in the task. However, as seen in Figure 2.11, the novice was significantly more quick (i.e., less careful) to determine that an attack occurred with less accumulated threat evidence in comparison to the expert (z = 3.816, p < .001). The researchers suggest that cyber security experts may have the tendency to be more cautious and conservative when determining that an attack occurred because of the high costs for wrongly classifying an alert in a IDS, however the response bias was not calculated in order to be certain.

Overall, their study merely tells us that the experts performed better than the novice in an IDS because of their cyber knowledge, and that they investigated the alerts more thoroughly than the novice before making a determination of an attack. Besides knowledge which was not always the case, it does not tell us why experts and novice differ. The analysts further differ in their cognitive abilities behind their investigative behaviors that contribute to a more successful performance, although not identified in the study. Further correlational research is required in order to investigate which cognitive factors/abilities most associate with knowledge (Motowildo et al., 1997) and to the performance in an IDS task.

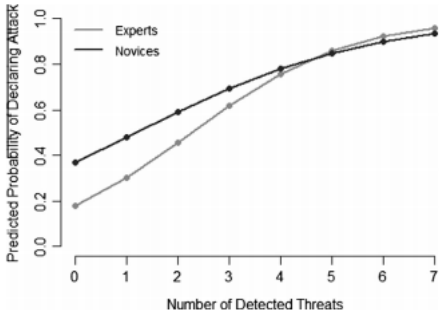


Figure 2.11. The Figure Shows the Probability of Experts and Novices to Declare a Cyber-attack Depending on the Number of Correctly Detected Threats, and Indicates that Novices Declare a Cyber-attack Quicker and with Less Evidence (Ben-Asher & Gonzalez, 2015).

Further, the study highlights the difficulty of acquiring individuals with actual IDS expertise to participate in laboratory studies. Here, like in many of the cyber performance studies a student group without any experience at all in network security/IDS work is pooled as a comparison group (i.e., novice) and asked to complete a “real-world” IDS task when it takes years of expertise to operate one (Goodall et al., 2009a). That drastically reduces its ecological application. Lastly, although many of the methods under SDT was applied in the study which offers more precise performance differences, the response bias and/or the decision strategy under SDT was not reported.

2.4.4 Response Bias

The response bias metric (i.e., decision strategy) under SDT is defined as, “reflecting the outcome of a decision making process that occurs as an individual is faced with choosing between two or more options. Response bias quantifies the tendency to either respond in a predominantly liberal (i.e., yes-threat) or conservative (i.e., no-threat) direction” (Waring, Chong, Wolk, & Budson, 2008, p.13). Besides being a performance measure in experimental research where it is manipulated and then compared, the response bias has also been seen as a possible trait characteristic using the individual differences framework. Evidence of this first comes from the research (Klauer & Kellen, 2010; Van Zandt, 2000), finding that individuals vary from a liberal bias to a conservative and/or cautious bias in most tasks. Additionally, the response bias has been seen as reliable and stable across and within recognition memory tasks, offering evidence that that the response bias is trait-like. For instance Kantner and Lindsay (2012) showed that a liberal bias in their first memory experiment remained a liberal bias in the second (Kantner & Lindsay, 2012, 2014). Further, the response bias from memory tasks correlate with other stable cognitive abilities such as executive function (e.g., working memory and prepotent response inhibition) (Huh, Kramer, Gazzaley, & Delis, 2006; Kantner & Lindsay, 2012, 2014), and with personality traits (Gillespie & Eysenck, 1980). In short, the findings from the correlations are said to mean that individuals with a more liberal bias tend to accept memories as true often with little memory evidence (Kantner & Lindsay, 2012). Overall, the response bias resembles a stable cognitive trait where individuals differ in the amount of evidence they require to make decisions (Kantner & Lindsay, 2014).

When the response bias is applied to an incident detection task (Ben-Asher & Gonzalez, 2015; Dutt et al., 2013; Rajivan et al., 2013), some cyber analysts will be more inclined to judge alerts as non-threatening, and will therefore have a bias to say no-threat. Other analysts may tend to judge the alerts more suspiciously, regardless of the packet contents and/or collected evidence. Those participants have a bias to say yes-threat. The human factors research in cyber defense has very minimally investigated the decision making strategy of analysts in regards to their response bias.

Dutt et al. (2013) first applied a computational cognitive model (under CyberSA), to expected cyber defender behavior in various attack and defend situations. The models predicted that decisions in an IDS depends upon the cyber analysts subjective risk-level (i.e., tolerance), or willingness to classify an alert as a threat. In other words, a low tolerance to threats was predicted for better accuracy in detecting threats (Dutt et al., 2013), however real behavioral data was not used. That prediction is representative of a liberal response bias under SDT; meaning to obtain a higher false-alarm rate but less misses. That response is safer than wrongly classifying a true-positive threat as benign (Ben-Asher & Gonzalez, 2015). One other study (Rajivan et al., 2013) calculated a conservative bias for two groups of cyber analysts however, the participants were not actually cyber analysts with IDS experience. Instead they were recruited from the university subject pool and trained to perform the task. Additionally, the calculation for the response bias was incorrect.

When a main goal in cyber defense work is to avoid wrongly classifying true-positive threats as benign because of the high costs in doing so, the accuracy score measuring percent correct only may be an inadequate measure for cyber performance, as it does not measure the costly mistake by itself (D. M. Green et al., 1966; Lynn & Barrett, 2014). Additionally, when benefits and costs (i.e., pay-offs) differ like it does in a network environment it needs to be accounted for in the accuracy score (Lynn & Barrett, 2014) through calculating the individuals response bias. For now it remains unclear how individual differences in cognitive ability contributes to the response bias as a measure of cyber performance.

2.4.5 Signal Detection Theory Framework Summary

The cyber human-factors research was able to show a specific behavior that differentiate the more successful analyst from the less successful, that can be further explained by their individual differences (Sauce & Matzel, 2013). That behavior difference between the novice and the expert cyber analyst pertained to their "response bias" or in other words, how much evidence the analyst requires before making a threat decision (Ben-Asher & Gonzalez, 2015). Since knowledge did not appear to contribute to evidence accumulation in the cyber task from the study by Ben-Asher and Gonzalez (2015), the cognitive ability behind that behavior when actual cyber tools are provided may contribute more to its success. Therefore, it remains unclear how information search and evidence accumulation, the main cognitive task for identifying threats in triage analysis work (D'Amico et al., 2005) depends on the analysts cognitive abilities.

In most of the studies presented thus far, the cyber performance scores from a variety of cyber tasks has shown as consistently poor (Ben-Asher & Gonzalez, 2015; Champion et al., 2012). That suggests environmental factors that are not under the control of the analyst is affecting performance. The SDT methods can offer a more complete picture of cyber performance differences and it can further show how the individual analyst is affected by the ambiguous/conflicting threat information in an IDS. For instance, the response bias (N. Macmillan & Creelman, 2005) accounts for the decisions that are made under that ambiguity (i.e., decision strategy) but it is rarely accounted for in cyber performance studies. Discussed next is one last theoretical framework of cognitive processes/concepts for cyber defense called the

Hybrid Space (Jøsok et al., 2019). It has been used to indirectly measure the cognitive agility of individual cyber analysts- using a between-subject approach.

2.5 The Hybrid Space Framework

The Hybrid Space is a conceptual framework (Jøsok et al., 2019) that was designed to represent a cyber operators range of cognition (i.e., cognitive agility) when conducting cyber defense operations. An analyst is asked to mark their current work focus in one of the quadrants in the Hybrid Space diagram, shown in Figure, 2.12 as they perform their real-time cyber analysis.

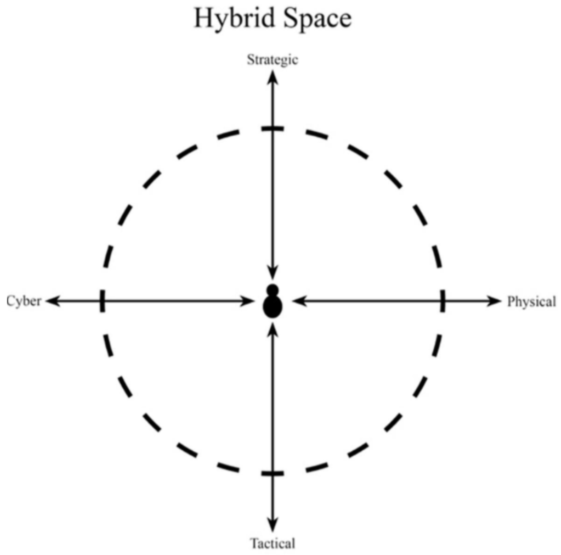


Figure 2.12. The Hybrid Space Diagram (Jøsok et al., 2019; Knox et al., 2017)

The movement (total distance traveled, x and y movement and quadrant changes) in the Hybrid Space is considered a measure of cognitive agility. The following human factors studies are among the few that seek to identify individual differences in cognitive ability, and how it contributes to cyber performance.

2.5.1 Cognitive Agility and Self-Regulation

Individual differences in an aspect of attention known as cognitive agility, and the self-regulation ability has been proposed to underlie cyber performance (Jøsok et al., 2019;

Knox et al., 2017). The researchers (Jøsok et al., 2019; Knox et al., 2017) using the Hybrid Space defined cognitive agility with following three cognitive concepts (Good & Yeganeh, 2012).

- Cognitive flexibility: ability to shift mental sets and adjust behavior to a changing environment
- Cognitive Openness: being open to new perspectives, ideas and experience
- Focused Attention: ability to select the relevant targets from the irrelevant targets and to also ignore distracting information

Cognitive agility in cyber defense is referred to as, the response or adjustment to suddenly changing environmental demands (Knox et al., 2017). Knox et al. (2017), describes it as,"cognitive focus movements" in the Hybrid Space Framework. In regards to the self-regulation concept, it also has many definitions and theories attached to it, but all of them are concerned with an individuals control (i.e., regulation) over their actions, thoughts, emotions, impulses or appetites, and task performances (Banfield, Wyland, Macrae, Münte, & Heatherton, 2004; Miyake & Friedman, 2012). Self-regulation can be defined as, "The higher order (i.e., executive) control of lower order processes responsible for the planning and execution of behavior" (Banfield et al., 2004, p.62). The cyber performance research (Jøsok et al., 2019; Knox et al., 2017) using self-regulation in cyber performance studies measured it through subjective trait questionnaires, but it can also be measured through direct measures of cognitive ability (Miyake & Friedman, 2012), as follows.

The 3-factor model of executive function is the most widely studied executive function (EF) model for individual differences research (Karr et al., 2018). The three proposed functions and or constructs of the central executive include:

- Updating
- Inhibition
- Set-shifting

The updating construct includes maintaining relevant information to working memory. A higher working memory may allow for a quicker transition to a new task (i.e., flexibility) (Kiesel et al., 2010; Pettigrew & Martin, 2016). Further, individuals with higher working memory may experience less distractibility (Lecerf & Roulin, 2009); and/or apply better inhibition (i.e., attention control) during set-shifting (Engle & Kane, 2004; Kane & Engle, 2003). The inhibition construct includes inhibiting information that is not task-relevant. It is summoned during set-shifting tasks to suppress or inhibit competing task-sets and distracting information (i.e., interference) (A. Allport & Wylie, 2000; D. A. Allport et al., 1994) so that attention can focus on and/or select the relevant information for the task (Engle & Kane, 2004). Individuals vary in their ability to inhibit the distracting information that appears during task-switching (i.e., a top-down task) (Engle & Kane, 2004; Miyake & Friedman, 2012), and that represents their executive attention control. Lastly, the set-shifting construct refers to the time that it takes to shift attention (i.e., ease of transitioning) to different tasks (e.g., rules and task-sets) (Miyake & Friedman, 2012; Miyake et al., 2000).

Cyber defense researchers (Jøsok et al., 2019; Knox et al., 2017) have used the Hybrid Space Framework to measure cognitive agility (i.e., self-reported cognitive location) in association with a trait questionnaire of self-regulation, both of which were measured over a four day cyber defense exercise. The exercise required 23 cadets with experience to defend a network from attacks. Through the use of a linear regression model, all self-regulation variables predicted Hybrid Space movements (cognitive agility), seen in Figure 2.13.

The higher levels of self-regulation explained 43.1 percent (alpha was set to 0.05) of the total cognitive movements in the Hybrid Space. The researchers suggest that higher levels of self-regulation was associated with a more ‘active search’ for relevant information (Jøsok et al., 2019). However, there are a few notable issues with the study.

The researchers stated that, “validating self-regulation as a contributing factor to cognitive agility is important as this can be a pathway to empirically underpin individual cyber operator performance” (Jøsok et al., 2019, p.1). First, self-regulation and cognitive flexibility is already a well-established construct (Miyake & Friedman, 2012). The executive functions are a core component of self-regulation ability and to flexibility (Miyake & Friedman, 2012) that can be easily measured with a traditional set-shifting task. Second, although identifying the cognitive

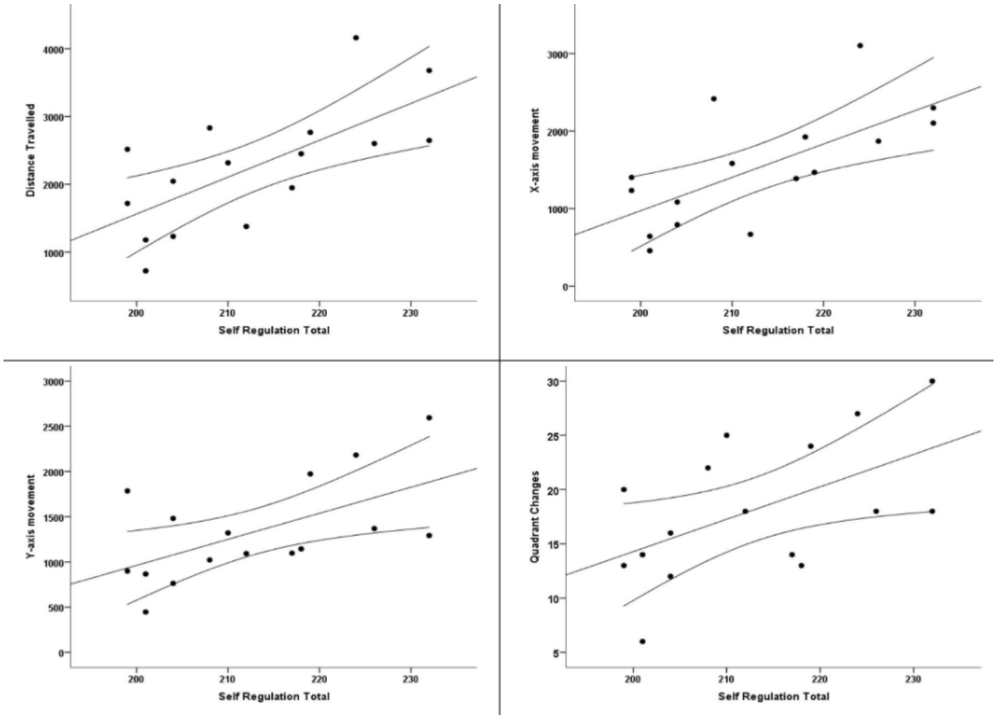


Figure 2.13. Scatter Plots Comparing Self-regulation Variables and Cognitive Movements in the Hybrid Space (Jøsok et al., 2019)

ability that mostly underlies cyber performance is important, it is more scientifically appropriate to identify them naturally rather than forcing self-regulation to basically correlate with itself in another measure of subjective "cyber flexibility" (Cronbach, 1957). Furthermore, (Robertson & Kandola, 1982, p.11) warned that, "researchers should not attempt to increase validity by developing criteria that are likely to relate closely to the predictor. Rather, it is important that the criteria themselves are job performance measures". Overall, further research with reliable and direct cognitive measures are needed to confirm the associations between self-regulation, agility and and cyber performance. Also, the cyber performance can be better measured in a more controlled, simulation/environment rather than in a cyber defense exercise. Significant associations would then suggest that flexibility is a core requirement for an analysts' cognitive movement through cyber tasks.

2.5.2 Hybrid Space Summary

The research mentioned in the section that applied subjective measures and/or questionnaires find indirect associations between cyber performance and higher level cognitive ability. Although useful in detecting individual differences, the method does not allow for direct identification of the cognitive abilities that are driving cyber performance scores. Furthermore, the cyber task used in a majority of cyber performance research does not represent any specific cyber defense work-task. In order to properly identify the individual differences that contribute to cyber defense success, a specific real-world cyber task/tool must be used under normal conditions (i.e., no experimental manipulations) (Sauce & Matzel, 2013), and data must be collected from actual analysts with experience using that tool.

2.6 Information Search and Evidence Accumulation Behavior Differences

A key behavior difference that has been identified in cyber performance studies between the novice and the expert cyber analyst is their information search behavior into security events through cyber tools, with the expert showing more evidence accumulation prior to making threat decisions (Ben-Asher & Gonzalez, 2015). Further, (Jøsok et al., 2019) found higher levels of self-regulation in association with a more 'active search' for relevant information (Jøsok et al., 2019). Lastly, a series of conference studies from Sandia National Laboratories (Emmanuel et al., 2015; Silva et al., 2014; Stevens-Adams et al., 2013) collected human-machine transactions of analysts (n = 26) through automated data logging, as they participated in a cyber training exercise. They found that participants using more tools (e.g., Wireshard, Reverse Engineering Binary, and Reverse Engineering Java) over the course of the training, may have performed better than those using less, as seen in Figure 2.14.

In a follow-up to the study (Silva et al., 2014), the cyber expert was then identified as being more thorough in their investigations in comparison to less experienced analyst, as discovered from their higher tool-use. For instance, a high positive correlation between cyber experience and specific tool-use (r = 0.565) was discovered. The higher tool-use behavior of the

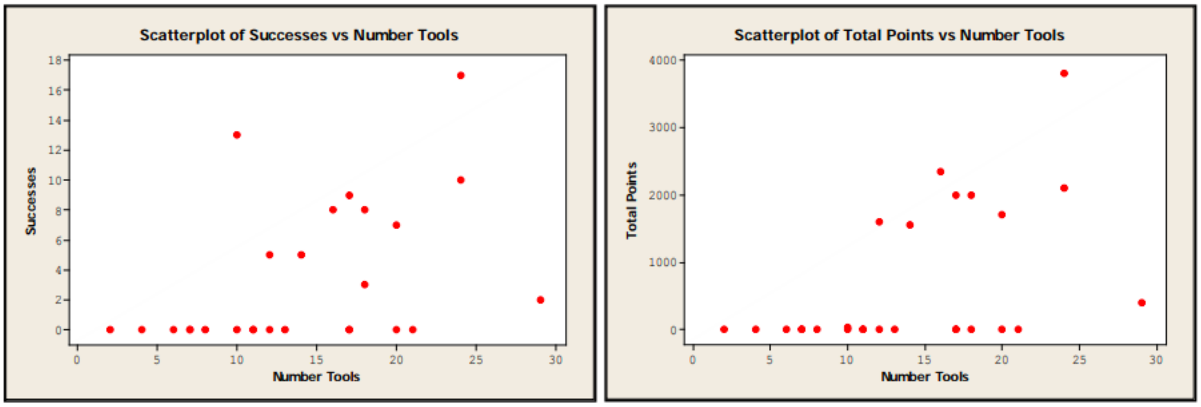


Figure 2.14. Scatter Plots Comparing Accuracy (i.e., Number of Successful Answers Submitted) and Tool-use (i.e., Number of tools used) (Stevens-Adams et al., 2013)

expert was contributed to a better task understanding for the expert that allowed them to more freely accomplish goals with a variety of different tools.

Although there are various method issues with the conference contributions, such as the use of non-standardized performance measures and non-representative cyber tasks, the research does pinpoint a specific behavior, rather than performance results, that differentiate the more successful analyst from the less successful in cyber defense tasks that can be further explained by their individual differences (Motowildo et al., 1997; Sauce & Matzel, 2013). Individual differences in cognitive ability could be driving the search performance difference.

2.7 Chapter Summary

The human-factors research in cyber defense has provided cognitive factors that affect cyber performance and specific behaviors that differentiate the more successful analyst from the less successful in cyber defense tasks that can be further explained by their individual differences (Sauce & Matzel, 2013). The identified behaviors and factors in IDS work serve as clues for the current individual differences study towards identifying the cognitive abilities that mostly underlie cyber performance. The identified cognitive factors that influence cyber performance included cognitive data-load (Champion et al., 2012; Dutt et al., 2013; Greenlee et al., 2016), vigilance (i.e., the vigilance decrement) (McIntire et al., 2013; Sawyer et al., 2015), spatial attention (Greenlee et al., 2016), task-switching (Vieane et al., 2017, 2016), flexibility (Knox et al., 2017) and knowledge (Ben-Asher & Gonzalez, 2015). Here, there are unidentified individual differences in cognitive ability that can mitigate the negative effects that certain factors have on cyber performance (Motowildo et al., 1997; Sauce & Matzel, 2013). The identified human-computer behavior differences included investigatory search differences, identified through higher tool-use for experts (Ben-Asher & Gonzalez, 2015; Emmanuel et al., 2015; Jøsok et al., 2019), and response bias (i.e., decision strategy) (Ben-Asher & Gonzalez, 2015; Dutt et al., 2013; Rajivan et al., 2013). Again, there are unidentified individual differences in cognitive abilities behind the investigative behaviors and decision strategy that contribute to the analysts' cyber performance.

Overall, the human factors research has not yet identified which cognitive abilities are contributing most to cyber performance (i.e., information search and evidence collection) and there is very little research seeking to measure the individual differences (i.e., between-subject variance) in basic human cognition (i.e., abilities) and how much they contribute to cyber performance success. Furthermore, there are significant methodology issues in the human-factors research aiming to study cyber performance. The main issues seen are that researchers use college students rather than actual analysts with cyber expertise. They use make-shift cyber exercises or tools that have not been validated in previous research, and further do not represent real-world cyber tools that are specific to a cyber work-role. Last but not least, researchers attempt to increase validity by developing criteria that are likely to relate closely to the predictor Jøsok et al. (2019); Knox et al. (2017); McIntire et al. (2013); Sawyer et al. (2015). Finally, precise measures of performance are not always applied.

CHAPTER 3. FRAMEWORK AND METHODOLOGY

The following chapter provides the methodology and framework that was applied in the research study. First an overview of the entire study is provided, followed by detailed discussion of each cognitive ability measure used in the study. Each cognitive ability variable is subdivided into three subsections in order to discuss the related stimuli, procedure and analysis. Following this, the variables of cyber performance are covered similarly. A description of the study's sample, along with the hypothesis being tested, and what determines a successful test for rejecting the null hypotheses will be discussed. The rationale for applying the measures is given throughout the section.

3.1 Overview of the Study

The research question for the study was: How do individual differences in attention control contribute to the cyber performance of network security analysts? The hypothesis involved investigating the associations between cognitive ability measures (the predictor variables) and cyber performance scores (criterion variable) and whether or not the cognitive ability can predict performance. In order to answer the research question, first a measure of cyber performance was obtained in a real-world cyber task under normal conditions. The human-computer interaction from 16 cyber analysts from Purdue University with expertise in using an incident detection system (IDS) was collected while they performed in a IDS simulator known as the Cyber Intruder Alert Testbed (CIAT) (Funke et al., 2016). The IDS task required the analyst to freely investigate, and to then classify 40 alerts that continuously arrived in a dashboard as either threat or no-threat (i.e., a binary classification). The CIAT simulator has been applied in cyber defense human factors research various times (Borneman, 2018; Funke et al., 2016; Greenlee et al., 2016; Vieane et al., 2017, 2016) with great results. The study then assessed the human computer interactions under Signal Detection methods (N. Macmillan & Creelman, 2005; Stanislaw & Todorov, 1999) in order to obtain a precise and complete picture of the analysts decision-making performance in the IDS environment. The SDT calculations provided a hit-rate and a false-alarm rate from the accuracy score.

Next, to properly assess the variation in cognitive ability (i.e., attention control) and its impact on cyber performance, the cognitive ability data was collected through direct measures of attention and executive function. The direct measures allowed for a more precise measure of cognitive ability, the closest to measuring actual neural activity as possible. The cognitive ability measures were selected after a full review of the human factors that were seen to influence cyber performance. Only the cognitive ability measures designed for individual differences research was selected for the current study. First, the Attention Network Task (ANT) (Fan et al., 2005, 2002) was selected for the current study, which measures individual differences in spatial-attention through the efficiency of the alerting (i.e, vigilance), orienting, and executive control networks (Engle & Kane, 2004; Miyake et al., 2000; Posner & DiGirolamo, 1998). Next, a digit-span was applied to assess working memory capacity (Baddeley & Hitch, 1974), and finally a task-switching measure (Armbruster, Ueltzh, Basten, & Fiebach, 2012; Miyake & Friedman, 2012) was used to assess individual differences in flexibility under various conditions including task-switching under conflict. Those traditional measures of attention control have predictive validity in that they produce significant associations with performance in dynamic domains and they have been extensively studied. Thus, the tasks have good construct validity and their reliability is known.

Overall, studying the associations between the cognitive ability measures and cyber performance scores revealed the underlying processes or mechanisms mostly required to make more successful threat decisions in an incident detection system and answered the hypothesis for the current study.

3.2 Cognitive Ability Measures

The following section individually discusses the three cognitive ability measures that serve as predictor variables and/or as measures in association with cyber performance (i.e., criterion) (a largely cognitive criterion) and the reasoning for their selection is provided. In addition, the Cronbach reliability scores are provided per cognitive measure because the reliability of the measures determine the strength of the hypothesised correlations in the current study (Cooper, Gonthier, Barch, & Braver, 2017).

The within-subject reliability for cognitive measures is known to be on the lower side despite their common use (Hedge, Powell, & Sumner, 2018; Meyerhoff & Papenmeier, 2020), with the standard of good/substantial reliability for cognitive measures at (Cronbach's Alpha = 0.60) (Hedge et al., 2018). However, (Nunnally, 1978) strictly defined a reliability coefficient of 0.70 as modest reliability, and that value is considered acceptable for early stages of research; while the reliability of 0.80 is considered adequate for basic research. Therefore, the current study applied cognitive measures with reliability above (Cronbach's Alpha = 0.70), although some of the conditions within the measures were found to be lower. The consequences of that are seen in the analysis.

All of the cognitive measures applied in the current study were administered through a computer running Presentation software by Neurobehavioral Systems. Presentation software is a stimulus delivery and experiment control program for neuroscience that is designed to provide precise stimulus delivery and accurate response logging. The Attention Network Test (Fan et al., 2002) is discussed first.

3.3 Attention Networks

The Attention Network Test (ANT) (Fan et al., 2005, 2002), is a quick and popular spatial-cuing task that was developed to measure the efficiency of visual attention networks (Posner, 1990), and for use in individual differences studies. The ANT measure requires participants to perform a typical on-going Flanker task (Eriksen, 1995) among three conditions (i.e., executive control, orienting and alerting). How efficiently individuals orient their attention with various top-down verses bottom-up cues (i.e., priming) to the target in the Flanker task (Enns & Richards, 1997; Posner, 1990) is investigated. In most studies using the ANT, the larger orienting and alerting effects is said to reflect poor selective attention (Broadbent, Cooper, FitzGerald, & Parkes, 1982).

The various research (Fan et al., 2002; Ishigami & Klein, 2010; MacLeod et al., 2010; Rey-Mermet, Gade, Souza, von Bastian, & Oberauer, 2019) measuring the reliability of the ANT finds that it is a reliable and valid measure of spatial attention. Starting with the original study by (Fan et al., 2002), the executive control network was seen as the most reliable with an adequate test-retest score (Cronbach's Alpha = 0.77). The orienting network had a intermediate reliability score (Cronbach's Alpha = 0.61), and the alerting network was found to have the least reliable test–retest score (Cronbach's Alpha = 0.52) (Fan et al., 2002). The reliability of the executive control network is said to be higher because it is more directly measured (Fan et al., 2002; Ishigami & Klein, 2010), while the lower reliability for the alerting and orienting network may be because of their interaction (Ishigami & Klein, 2010). Other studies (MacLeod et al., 2010) found nearly the same split-half reliability results with 1,141 participants. Newer research (S. Campbell, 2016) exploring the reliability of the ANT, found a higher split-half reliability correlation score that ranges from (Cronbach's Alpha = 0.67) to (Cronbach's Alpha = .81) and finds that it remains reliable and consistent over one year. To date, those numbers are said to represent one of the best measures of attention network efficiency available.

The ANT measure was chosen for the current study because of its reliability and construct validity (Posner, 1990), and also because it has transferred from a laboratory to real-world situations/domains, such as the traffic and transportation domain (Roca, Crundall, Moreno-Ríos, Castro, & Lupiáñez, 2013; Weaver, Bédard, McAuliffe, & Parkkari, 2009) which found significant correlations between the attention networks and driving performance. Further, the musical domain (Medina & Barraza, 2019) found through the ANT that professional musicians outperformed non-musicians in their ability to resist the distracting Flankers in the executive control component to suggest better attention control for musicians. One last reason the ANT task was chosen for the current study is because vigilance was identified in the human factors research as influencing cyber performance (Greenlee et al., 2016; McIntire et al., 2013; Sawyer et al., 2015). The current study hypothesised that the attention networks would contribute/correlate highly with cyber performance.

3.3.1 Stimuli

The same version of the ANT described by (Fan et al., 2002) was applied in the current study. The main task of the ANT is a classic Flanker task (Eriksen, 1995) that can be seen in Figure, 3.1, where a row of five black arrows is displayed. The central arrow is set as the target, while the rest of the arrows are designated as flankers. The participants were required to respond by pushing a response button when the target arrow was facing left, and to push the button when the target arrow was facing right.

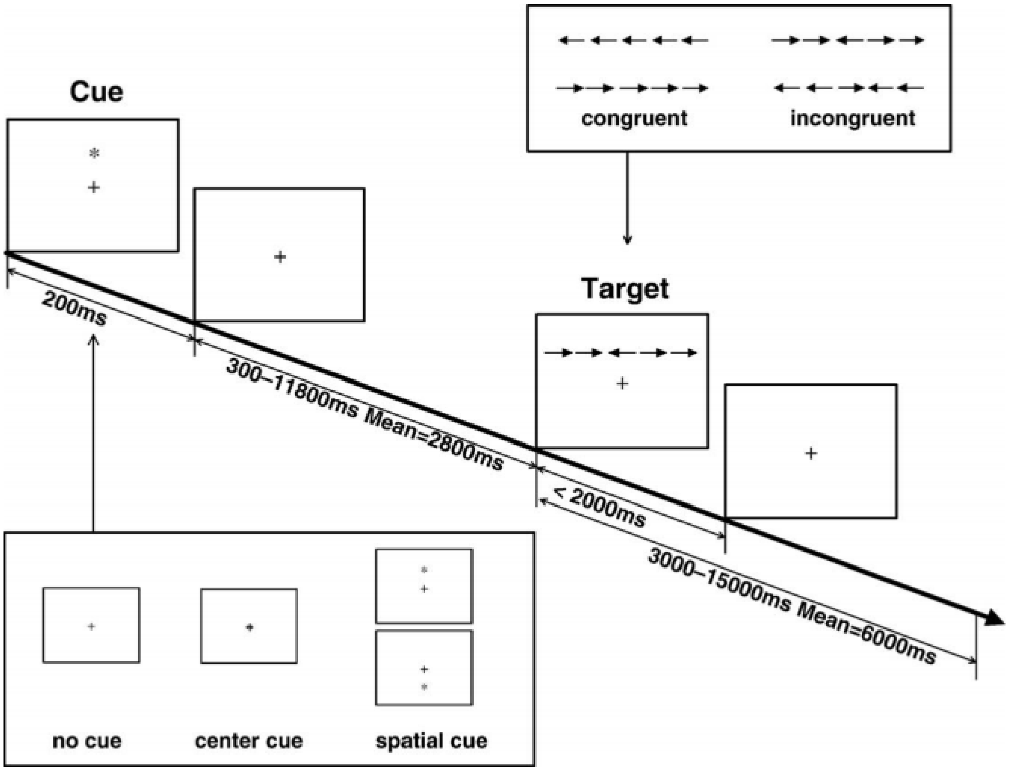


Figure 3.1. The Attention Network Task (Fan et al., 2005)

To summon executive control (i.e., conflict) in the task, congruent and incongruent stimuli was used. For the congruent condition, the flankers on each side of the target arrow faced the same way as the target arrow. For the incongruent condition, the flankers on each side of the target arrow faced the opposite direction. For the orienting component, the row of arrows was presented either above or below a fixation point, at random. That required participants to quickly shift/orient their spatial attention up or down. For the alerting component, three cue conditions were used (no cue, cue, center-cue). An asterisk served as the cue to where and when the row of black arrows would appear on the screen. The center-cue condition alerted the individual for when the the black arrows would appear. The spatial-cue condition alerted participants of when and also where the black arrows would appear on the screen. Lastly, there was a condition in which no cue was given (i.e., no-cue condition). That condition gave an individuals baseline of sustained attention only (Fan et al., 2005, 2002).

3.3.2 Procedure

The attention network task (Fan et al., 2002), was administered through a computer running Presentation software by Neurobehavioral Systems. Each 17 minute session consisted of one 2 minute practice block with performance feedback, and three experimental blocks without feedback. The experimental blocks each took five minutes to complete, and consisted of 96 trials (four cue conditions x two target locations x two target directions x three flanker conditions x two repetitions). The participants gave their responses (i.e., finger-presses) through a response box that was connected to the computer running Presentation software, where the response time was precisely recorded. The long trials in the ANT aid in reducing the proportion of error variance (Nunnally, 1978) for better reliability in individual differences research.

3.3.3 Analysis

The three attention networks scores were calculated by the following subtractions. To calculate the Alerting Network contribution (presence or absence of cues without spatial information); the difference between trials that were preceded by a double cue (rt - dc) and those that had no cue (rt - nc) was computed. To calculate the Orienting Network contribution (presence or absence of cues with spatial information); the response difference between trials that were preceded by spatial cues (rt - sc) and those with central cues (rt - cc) was computed. To compute the Executive Network contribution; the differences between congruent flankers (rt - c) on both sides of the target and incongruent flankers (rt - i) was computed. The faster response scores indicate greater efficiency, for alerting and orienting networks. For the EF network, a lower score indicates greater attention control/efficiency.

The descriptive statistics for each network can be seen in Table, 3.1. The median RT scores for each network was taken for further analysis, as those values are less likely to be influenced by outlier trials. The average accuracy of ANT performance was high (0.95 and up) indicating that all participants understood the instructions and were able to perform the task reliably. The mean for the executive function network was (m = 77 ms, SD = 19). The orienting network mean was (m = 36 ms, SD = 13), and the alerting network mean was (m = 43 ms, SD = 22). A previous individual differences study (McConnell & Shore, 2011) using the same ANT task on 588 healthy older individuals found similar means and error in comparison to the current study, except that the executive control mean obtained here for the cyber analysts was a bit better. The bi-variate two tailed correlations between the three networks was addressed and there were no significant correlations between any of the attention networks. Overall, the data obtained from the rather small sample of cyber analysts appears accurate and has a wide range of scores.

Table 3.1. *Descriptive Statistics*

	Alerting	Orienting	EF Median
Valid	15	15	15
Missing	0	0	0
Mean	43.1	36.2	77.0
Std. Error of Mean	5.7	3.3	5.1
Median	41.2	34.5	80.9
Std. Deviation	22.2	12.9	19.7
Range	69.9	40.8	54.6
Minimum	10.1	16.4	48.7
Maximum	80.0	57.2	103.3

3.4 Digit-Span

The simple digit-span task (Woods et al., 2011) also known as the backwards and forwards digit-span was selected for the current study to measure individual differences in visual, short-term working memory capacity (Baddeley & Hitch, 1974). The forwards digit-span includes remembering the order of a perceptually observed sequence of numbers. The span mainly measures a individuals storage capacity of visual information, once perceptual information is removed (Chun, Golomb, & Turk-Browne, 2011), whereas the backwards-digit span includes observing a sequence of numbers and then repeating the sequence in reverse order. The backwards span measures an individuals ability to maintain the information in short-term storage as well as to internally manipulate that information (Shipstead et al., 2015). The manipulation draws more on attention/executive control (i.e., inhibition processes) (Engle & Kane, 2004; Fougnie & Marois, 2007).

The first reason for selecting the digit-span for the current study is because of its reliability estimate, often calculated in the 0.8 to 0.9 range in experimental research (Engle & Kane, 2004). The digit span chosen for the the current individual differences study (Woods et al., 2011) reported a “high” test-retest correlation (Cronbach’s Alpha = 0.68) for the forward span, in regards to the longest list correctly reported on any of the 14 trials. The test-retest correlations for the backwards span was actually high (Cronbach’s Alpha = 0.81) for the maximum list reported. The test-retest reliability was taken over three days (n = 31) with an age range of 18 to 46 (mean age = 26 years) with an average of 14.8 years of education. The high reliability for WM allows it to successfully predict higher-level cognitive abilities in individual differences research, such as reading comprehension (McVay & Kane, 2012), writing (Daneman & Carpenter, 1980), and general intelligence (Unsworth & Engle, 2007). The digit span is also used in standardized testing (Wechsler Intelligence Scales). The main and final reason for applying the WM digit span in the current study is because cognitive task analysis research (D’Amico et al., 2005), and cognitive decision models (Dutt et al., 2013) in cyber performance research posits working memory ability as a main contributor to the primary task of IDS work in the formation of mental models and data-fusion. The current study hypothesized that a higher working memory capacity may allow for more efficient cyber performance in IDS work.

3.4.1 Stimuli

Participants completed a computerized forwards and backwards digit span from the Poldrack Lab at Stanford University (Woods et al., 2011). First the instructions appeared on the computer screen for the forward digit-span, instructing the participant to try to remember a sequence of numbers that would appear on the screen one after the other. At the end of each trial, the participant was asked to enter all the numbers into a presented number-pad in the sequence in which they occurred, as seen in Figure, 3.2.

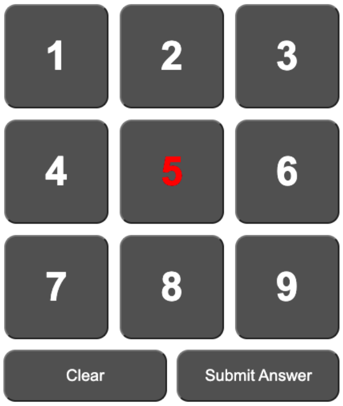


Figure 3.2. Poldrack Lab Digit-Span (Woods et al., 2011)

The participant used the computer mouse to click on the digits in the order in which they originally appeared. For the backwards digit-span the participant was asked to report the reverse of the sequence of numbers displayed.

3.4.2 Procedure

First the instructions for the forwards digit-span was displayed on the screen, and then the forward digit-span took place starting with a 3 digit length. Once the forward digit span was completed, the instructions for the backwards digit-span appeared on the screen, followed by the backwards digit-span starting at a two digit length. In both parts of the task, 14 trials were presented, with the digit length increasing by one digit after each correct trial and decreasing after two successive incorrect trials at the same list length.

3.4.3 Analysis

The working memory (WM) ability measure was scored for each participant by computing the highest number of remembered digits per span. That number indicated the participants' maximum working memory capacity. The descriptive statistics for the measure of WM is displayed in Table 3.2. The working memory measure applied to the sample of cyber analysts obtained a mean forward working memory span of (m = 8.3), and a mean backwards working memory span of (m = 7.5). Both scores are higher than the known population average (m = 7) (Miller, 1956). A study by Woods et al. (2011) that uses the exact test as the current study reported a mean for the forwards span at (m = 7.87) digits, and a mean for the backwards span at (m = 6.48) digits (n = 31). Therefore, the sample of cyber analysts in the current study had an elevated WM score, meaning they have a higher working memory capacity. The test still appeared to have enough variance to include it in the current study, with the backwards digit span having more.

Table 3.2. *Descriptive Statistics for the Working Memory Digit Span*

	WM Forwards	WM Backwards
Valid	16	16
Missing	0	0
Mean	8.31	7.50
Std. Error of Mean	0.33	0.35
Std. Deviation	1.30	1.41
Variance	1.70	2.00
Range	6.00	6.00
Minimum	6.00	4.00
Maximum	12.00	10.00

3.5 Task Switching

Flexibility was measured through a neurological measure called Stabflex (Armbruster et al., 2012). The task-switching measure was designed for use in individual differences research. The main on-going task for Stabflex is a typical on-going judgement task (i.e., magnitude and parity task) from the classic switching paradigm (Jersild, 1927; Rogers & Monsell, 1995). The Stabflex task involves switching between the magnitude and parity of a single digit- a task that is widely known to represent the speed and flexibility with which people switch between two different tasks. The Stabflex measure further adds three conditions that occur during the on-going task to include a task switching condition, a distractor inhibition condition, and an ambiguous condition. The ambiguous condition allowed for measuring conflict/choice during switching and/or how long it takes to react to the interruption, which is rarely addressed in usual set-shifting tasks.

The test-rest reliability of usual switching tasks have adequate reliability. For instance, (Pettigrew & Martin, 2016) reported a adequate internal (i.e., within one session) split-half reliability score (Cronbach's Alpha = 0.67) for switching tasks. The lower reliability estimates seen in set-shifting task are said to occur because the task may be too easy, thus causing low between-subject variability which then leads to low reliability for measuring individual differences (Hedge et al., 2018). However, a recent study (Kraft, Rademacher, Eckart, & Fiebach, 2020) using Stabflex (n = 100) reported a high internal (within subject) split-half reliability for the switch cost (RT) (Cronbach's Alpha = 0.87), but the switching error was inadequate (Cronbach's Alpha = 0.23), while the spontaneous switch rate (ambiguous condition) was also high (Cronbach's Alpha = 0.83). The Stabflex task was chosen for the current study because of its difficulty for which more use of attention control is required. Furthermore, it required extensive training for the task, which increases its reliability by reducing the proportion of error variance (Nunnally, 1978).

The predictive validity in the applied research for set-shifting tasks is growing. For instance, the updating of WM in a set-shifting task (Miyake et al., 2000) was able to predict an aviators flight performance (Causse, Dehais, & Pastor, 2011). In addition, the video game domain (Colzato, Van Leeuwen, Van Den Wildenberg, & Hommel, 2010; C. S. Green & Bavelier, 2006) has found set-shifting performance differences between individuals that play video games and those that do not. The video game players (VGP) were found with superior flexibility in comparison to non-video game players (Colzato et al., 2010). One last major reason for choosing the Stabflex task is because the human factors research in cyber defense identified task-switching behavior (Ben-Asher et al., 2015; Kortschot et al., 2018) and flexibility (Jøsok et al., 2019; Knox et al., 2017) as a possible highly contributing factor to cyber performance. The current study hypothesised that higher amounts of flexibility (i.e., lower costs) would correlate with more successful cyber performance.

3.5.1 Stimuli

The Stabflex task can be seen graphically in Figure 3.3. The response task was presented through a computer running Presentation software. A keyboard was required in order to give the responses. The participants assigned two fingers (middle and index finger) on each hand to four keyboard buttons that were mapped to odd and even on the left hand, and greater-than/less-than on the right hand. The stimuli for the task included digits one through nine in the color gray. The participants were told to always use the brighter of the two gray digits when performing the on-going magnitude/parity task.

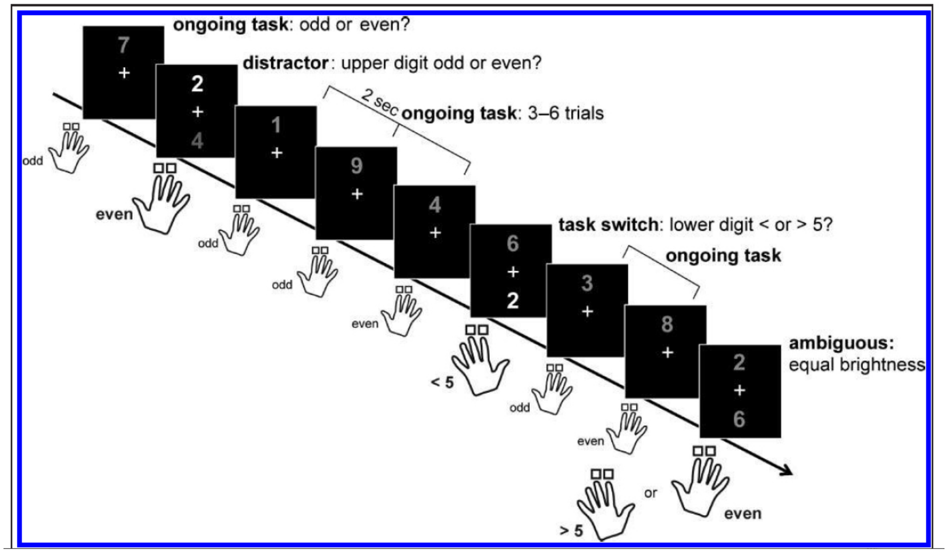


Figure 3.3. StabFlex Task
(Armbruster et al., 2012)

3.5.2 Procedure

Stabflex (Armbruster et al., 2012) was administered through a computer running Presentation software by Neurobehavioral Systems. First, the participants completed three training components. The first training component instructed the participant on how to complete the odd/even decision (i.e., on-going task) using one number that appeared above the fixation cross. That was followed by a practice run that included 20 trials with a number appearing above the fixation cross. In the second training component the participant was instructed on how to complete the smaller/larger than five decision that appeared below the fixation cross. That was followed by 20 trials with a number appearing below the fixation cross. At the end of the training, a percent correct was shown on the screen and if it was lower than 80 percent correct the participant repeated the training. Next, the participants completed one last training component with three blocks of training that incorporated both tasks from the prior two training. The first block included 60 trials with right/wrong feedback and no ambiguous trials. Block two included 80 trials with right/wrong feedback and no ambiguous trials and finally block three included 100 trials without feedback and with ambiguous trials included. At the end of the training, the percent correct was shown on the screen.

After successful completion of the training, the participants completed the five minute experimental task which was presented in two separate blocks with 150 trials each. The experimental task procedure included the baseline task (i.e., on-going odd or even task) which occurred between three conditions (i.e., critical trials); task switching, distractor inhibition, and ambiguous conditions. After every critical trial, participants continued to perform the baseline task for three to six trials. For the on-going task, only one digit was displayed above the fixation cross and participants decided if it was odd or even. That was considered the baseline condition. The error and RT from the condition was measured. The participants responded 240 times (80 percent). For the remaining 20 percent of trials, two digits were presented on the screen, above and below the fixation cross for the following conditions:

In the distractor condition (20 trials), a second unrelated digit in a dark color appeared below the fixation cross. The participants still attended to the usual brighter digit. The absolute number of errors and RT was totalled for a measure of performance in the condition. The distractor inhibition condition is said to stay low until participants arrived to it via on-going task. The distractor condition is used as a measure of WM stability.

In the switching condition (20 trials) the participants attended to the brighter digit, but the digit would switch to sometimes appearing below the fixation cross, where participants would then switch tasks to performing the less than or greater than five rule. Again, a RT and error rate was calculated as a measure of performance in the switching condition. The condition is used as a measure of WM flexibility.

Finally, the last condition within the on-going task was the ambiguous condition (20 trials). In the condition two digits appeared above and below the fixation cross, but they were both the same bright gray color. Thus, no cue existed and participants were free to switch or stay. In other words, the participant either continued with the task that they just performed (e.g., classifying the top digit as odd or even) in the on-going task or they could switch to performing the previously used task (less than or greater than five on the bottom digit). The switches in the condition is considered by researchers (Armbruster et al., 2012; Armbruster-Genç, Ueltzhöffer, & Fiebach, 2016) to be measuring the stability (i.e., depth) of an individuals attractor states or the stability of WM. A RT and error rate was again calculated as a performance score in the condition.

3.5.3 Analysis

The data from the main experiment (Armbruster et al., 2012) was used for data analysis. First, the error per task condition (i.e., error during distractor present, error during switching, error during on-going trial, error during the ambiguous condition) was calculated into a percent (error x 100/240.0). Similarly, the average mean response time (RT) was calculated for each task condition in milliseconds. Then the costs in behavioral performance was calculated by subtracting the mean RT and mean error rate in the baseline condition from the mean RT and mean error rate in the respective task condition (Armbruster-Genç et al., 2016; Monsell, 2003). Finally, the spontaneous switch rate was computed as the number of switches in ambiguous trials, divided by the number of ambiguous trials (20) (Kraft et al., 2020). For analyses of the data, all trials with an RT less than 150 ms was eliminated as suggested by previous research (Armbruster-Genç et al., 2016).

The descriptive data from all Stabflex conditions is consistent with previous research using Stabflex (Armbruster et al., 2012; Armbruster-Genç et al., 2016) on healthy adults. The raw descriptive statistics for each condition in Stabflex for the current study can be seen in Table, 3.4 and in Table, 3.3. The task-switching (RT) time for cyber analysts (m = 935 ms) was slightly quicker, but very similar to the data collected in previous studies using healthy adults (m = 960 ms) (Armbruster et al., 2012; Armbruster-Genç et al., 2016). Additionally, that condition and the distractor inhibition (m = 829 ms) from the current study both showed higher costs than the on-going task like it should (Armbruster et al., 2012; Armbruster-Genç et al., 2016). The descriptive statistics for the ambiguous condition (error-rate and response times) can be seen in Table, 3.3. The mean switching rate in the ambiguous condition was (m = 10.5 percent) and the amount correct when switching was (m = 7.3 percent). The mean time for the non-switchers in the condition was (m = 996.3 ms), and for those that switched the mean was (m = 1042.0 ms). There was a significant correlation between the ambiguous condition and switching RT, as typically seen in previous research using the Stabflex task (Armbruster et al., 2012; Armbruster-Genç et al., 2016; Kraft et al., 2020). All of which to suggest a valid measure of flexibility performance for the sample of cyber analysts. Lastly, the calculated costs for all conditions (ambiguous, distractor and switching condition) is displayed in Table 3.5.

Table 3.3. *Descriptive Statistics for the Ambiguous Condition in Stabflex*

	AmbiSwitch(Total)	AmbiSwitch(Correct)	AmbiNonSwitch(RT)	Ambiswitch(RT)
Valid	15	15	15	14
Missing	0	0	0	1
Mean	10.5	7.3	996.0	1042.0
Std. Error of Mean	1.4	1.0	63.5	72.2
Median	10.0	8.0	1076.1	999.8
Std. Deviation	5.6	4.0	245.8	270.3
Range	19.0	14.0	766.7	821.7
Minimum	0.0	0.0	580.3	705.4
Maximum	19.0	14.0	1347.0	1527.1

Table 3.4. *Descriptive Statistics for Stabflex Conditions*

	Ongoing(RT)	Distractor(RT)	Switch(RT)	Ongoing(Er)	Distractor(Er)	Switch(Er)
Valid	15	15	15	15	15	15
Missing	0	0	0	0	0	0
Mean	659.9	829.3	935.7	6.2	15.0	20.7
Std. Error of Mean	27.6	44.2	50.0	1.1	3.2	2.4
Median	646.7	799.1	856.5	6.3	10.0	20.0
Std. Deviation	106.7	171.3	193.5	4.3	12.5	9.4
Range	326.8	537.3	599.0	15.8	50.0	30.0
Minimum	526.9	601.4	646.8	1.7	5.0	5.0
Maximum	853.7	1138.7	1245.8	17.5	55.0	35.0

Table 3.5. *Descriptive Statistics for Costs*

	Switch Costs (RT)	Distractor Costs (RT)	Ambi-Switch Costs (RT)
Valid	15	15	14
Missing	1	1	2
Mean	275.76	162.33	159.79
Std. Error of Mean	36.52	24.25	27.62
Std. Deviation	141.43	93.93	103.34
Variance	20003.08	8823.26	10679.41
Range	544.26	349.37	390.00
Minimum	104.65	16.00	2.00
Maximum	648.91	365.37	392.00

3.6 Cognitive Ability Measures Summary

In individual differences research it is important to have a large variance in scores and low error, per measure to increase reliability (Cooper et al., 2017; Hedge et al., 2018). In the current study, the variance for the cognitive tasks appear large. In addition, the cognitive data collected from the sample is similar to previous samples of healthy adults, and therefore it appears to be valid.

3.7 The Cognitive Task: CIAT

The literature review identified the use of an incident detection system (IDS) as the primary task of the Incident Responder in the triage/incident detection role. Therefore, the current study selected the simulated network environment known as the Air Force Cyber Intruder Alert Testbed (CIAT) (Funke et al., 2016), to assess and define cyber performance (Motowildo et al., 1997). The CIAT was developed for human factors research in cyber defense (Funke et al., 2016), and it was designed to resemble industry standard incident detection software (IDS) such as ArcSight, AlienVault and IBM's Security Network Protection (XGS). That is important because many of the tasks used to study cyber performance are not representative to a real-world software or cyber defense work-task, which reduces and/or eliminates ecological validity. The IDS task used in the current study is specific to a cyber defense work-role so that the analysts' behaviors in the IDS can transfer to their real-world/job performance (Motowildo et al., 1997). In further support for the selection of the CIAT for the current study is that it has strong face validity from the various experts and analysts that have confirmed its representation to the standard IDS software (Funke et al., 2016), and to the incident detection work-role (Borneman, 2018; Funke et al., 2016; Greenlee et al., 2016; Vieane et al., 2017, 2016). It has also been used in various cyber defense performance studies (Borneman, 2018; Funke et al., 2016; Greenlee et al., 2016; Vieane et al., 2017, 2016) with successful results.

Although the researchers using the CIAT in those studies may have tweaked the simulator to fit their experiment, the base IDS task remained the same. For the current study, the CIAT was not manipulated, since the goal of individual differences research is to investigate the already existing variation between cyber analysts. As (Cronbach, 1957, p.671) said, “the correlator finds his interest in the already existing variation between individuals, social groups, and species”. One last benefit to using the CIAT simulator, and a primary goal of correlation research (Cronbach, 1957) is that it allowed the information search and evidence collection task to remain uniform for each participant contributing to a correlation, which reduced confounding variables from influencing the cyber performance scores.

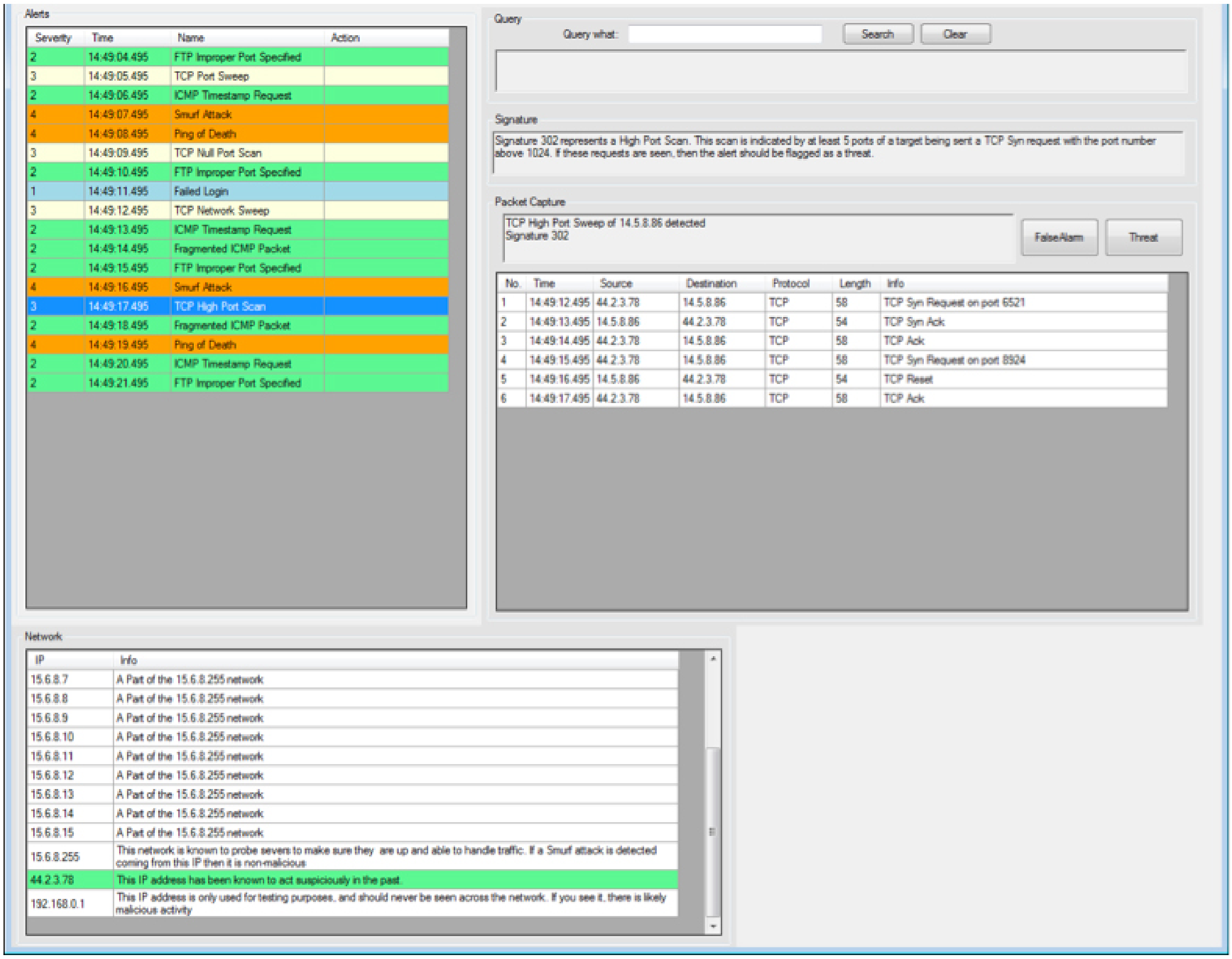
The CIAT task required each analyst to freely select individual alerts that appeared in a queue and to then decide if the alert was a threat, or no-threat. In order to make the classification decision various tools were available to the analyst that were located within tabs on the CIAT dashboard. The analyst used the tools to acquire more evidence in order to make the threat decision. In order for the alert to be a threat, all elements of a signature had to exist in the packet data and/or in the network log. Finally, it was hypothesised that cognitive ability (lower cognitive costs) would be associated with higher cyber performance accuracy, and with more efficient search behaviors in the CIAT.

3.7.1 Stimuli

The CIAT program was loaded on a Windows computer. Each participant in the study interacted with the original, unmodified CIAT program, which was pre-populated with data from scenario one, IP Set two (Funke et al., 2016). The CIAT dashboard can be viewed in Figure, 3.4. The dashboard was set-up to have a “multi-tab-view” where there was a tab for the alert queue (alerts tab), a signature database tab (query tab), a packet capture software tab (PCap tab) and a network list tab (network tab). A typical IDS will throw an alert when it detects network activity that is similar to a known signature. Likewise, the analysts used the alert tab (i.e., IDS) to classify the alerts/events in the queue as either a "threat" or "not a threat". The alerts in the queue were color coded by threat severity (e.g., severity levels one through five). The participants used the other three tabs (i.e., tools) to verify the existence of four to six pieces of evidence in order to classify the alert. For example, the query tab could be used to determine whether or not the alert matched the signature of a known threat, while the network tab displayed the associated network addresses, and the PCap tab displayed the associated packet data. An alert was confirmed as a threat if it matched all of the elements of the threat signature.

The original alert data-set provided with the CIAT program (Funke et al., 2016) was reduced from 90 to 40 alerts by removing the bottom 50 alerts. The decision to use only 40 alerts was made in order to fit the 45 minute time window for the task, and because the alert amount of 40 has been used in previous studies (Borneman, 2018; Greenlee et al., 2016; Vieane et al., 2017), using the CIAT simulator. The reduction left the alert severity distribution with the following: 12 percent of alerts were severity one, 30 percent of alerts were severity two, 27 percent of the alerts were severity three, 27 percent of alerts were severity four, and lastly two percent of alerts were severity five. The severity distribution can be compared to (Funke et al., 2016) which had; 25 percent of alerts that were classified as severity one, 25 percent were severity two, 20 percent were severity three, 20 percent were severity four, and 10 percent were severity five. The reduction of the data-set left the simulator with only two true-positive threats out of 40, while 38 alerts were false-negative threats. Thus, the base-rate (i.e., true-positive/false-positive ratio) became somewhat more representative to real world threat encounters, usually seen around one percent (Champion et al., 2012; Funke et al., 2016).

The first true-positive alert was located at alert number 23, and it was a high severity alert (category four). The alert type was a wiz attack (i.e., distributed denial of service) on the SMTP Server (pcap ID 9). The time stamp for that event was 12:30:30.455 (130000ms). The second true-positive alert was located at location 40. The alert was a high severity (category five) FTP heap corruption attack (pcap ID 36). The time stamp for the event was 12:32:30.455 (250000 ms). The rest of the 38 false alerts in the IDS are defined as, "to include instances where an IDS misidentifies normal, benign activity as malicious; detects only partial evidence matching a known signature; or correctly recognizes that malicious activity is present, but the system is not vulnerable to that activity (e.g., an attack exploits a vulnerability that has been patched, or an attack that targets a closed port)” (Funke et al., 2016, p.2). The alerts arrived in the IDS (i.e., alerts tab) in sets of 10, every two minutes.



Severity	Time	Name	Action
2	14:49:04.495	FTP Improper Port Specified	
3	14:49:05.495	TCP Port Sweep	
2	14:49:06.495	ICMP Timestamp Request	
4	14:49:07.495	Smurf Attack	
4	14:49:08.495	Ping of Death	
3	14:49:09.495	TCP Null Port Scan	
2	14:49:10.495	FTP Improper Port Specified	
1	14:49:11.495	Failed Login	
3	14:49:12.495	TCP Network Sweep	
2	14:49:13.495	ICMP Timestamp Request	
2	14:49:14.495	Fragmented ICMP Packet	
2	14:49:15.495	FTP Improper Port Specified	
4	14:49:16.495	Smurf Attack	
3	14:49:17.495	TCP High Port Scan	
2	14:49:18.495	Fragmented ICMP Packet	
4	14:49:19.495	Ping of Death	
2	14:49:20.495	ICMP Timestamp Request	
2	14:49:21.495	FTP Improper Port Specified	

Signature 302 represents a High Port Scan. This scan is indicated by at least 5 ports of a target being sent a TCP Syn request with the port number above 1024. If these requests are seen, then the alert should be flagged as a threat.

TCP High Port Sweep of 14.5.8.86 detected
Signature 302

No.	Time	Source	Destination	Protocol	Length	Info
1	14:49:12.495	44.2.3.78	14.5.8.86	TCP	58	TCP Syn Request on port 6521
2	14:49:13.495	14.5.8.86	44.2.3.78	TCP	54	TCP Syn Ack
3	14:49:14.495	44.2.3.78	14.5.8.86	TCP	58	TCP Ack
4	14:49:15.495	44.2.3.78	14.5.8.86	TCP	58	TCP Syn Request on port 8924
5	14:49:16.495	14.5.8.86	44.2.3.78	TCP	54	TCP Reset
6	14:49:17.495	44.2.3.78	14.5.8.86	TCP	58	TCP Ack

IP	Info
15.6.8.7	A Part of the 15.6.8.255 network
15.6.8.8	A Part of the 15.6.8.255 network
15.6.8.9	A Part of the 15.6.8.255 network
15.6.8.10	A Part of the 15.6.8.255 network
15.6.8.11	A Part of the 15.6.8.255 network
15.6.8.12	A Part of the 15.6.8.255 network
15.6.8.13	A Part of the 15.6.8.255 network
15.6.8.14	A Part of the 15.6.8.255 network
15.6.8.15	A Part of the 15.6.8.255 network
15.6.8.255	This network is known to probe severs to make sure they are up and able to handle traffic. If a Smurf attack is detected coming from this IP then it is non-malicious
44.2.3.78	This IP address has been known to act suspiciously in the past.
192.168.0.1	This IP address is only used for testing purposes, and should never be seen across the network. If you see it, there is likely malicious activity

Figure 3.4. CIAT Dashboard (Funke et al., 2016)

3.7.2 Procedure

First, the researcher demonstrated to the participant how to collect evidence in the IDS in order to decide if an alert was a threat or no-threat. For that, the researcher used the mouse to select the first alert in the queue and then instructed the analyst to read the data inside the alert, and to use the provided tabs (i.e., tools) to investigate the alert. The researcher then demonstrated how to classify the alert by clicking on the threat or no-threat button in the alert tab. The participant was given five minutes of practice time to become familiar with the network environment and IDS simulator. The participant was then given unlimited time to classify all 40 alerts in any order, but they were asked to perform at a work-pace without distractions. For each participant, the mouse and keyboard interactions with the CIAT software was logged and stored within the same computer.

3.7.3 Analysis

Traditionally, job performance was measured through either subjective methods such as supervisory performance ratings or through the work sample (i.e. performance test) (Hunter, 1983). A work sample directly measures job skill in a situation that is equal to that at work, under realistic and standardized conditions (Hunter, 1983; Motowildo et al., 1997). Work-samples are now considered among the most valid predictors of job performance (Hunter, 1983; Schmidt & Hunter, 2004). The CIAT simulator served as the work sample in the current study. It allowed for the collection of the following measures of cyber performance and investigator behavior.

This being an exploratory study, several measures of direct cyber performance were used to find out which one was most representative to the IDS task (Sokolova & Lapalme, 2009). First, the accuracy score and RT for alert classifications was collected and analyzed. Accuracy and RT are common metrics for the evaluation of classification models (Sokolova & Lapalme, 2009) for evaluating cyber defense performance (Ben-Asher & Gonzalez, 2015; Dutt et al., 2013; Funke et al., 2016; Vieane et al., 2016), and they are commonly measured in cognitive tasks (Lynn & Barrett, 2014). First, the total alert RT was taken because there is more variability in RTs than in accuracy rates, therefore they tend to have a higher reliability for studying individual differences (Cooper et al., 2017). Next, the overall accuracy was addressed (Accuracy = TP + TN/ TP + FP + FN + TN x 100 percent) (Sokolova & Lapalme, 2009). Accuracy measures overall performance and therefore encompasses all of the classification performances to include, the hit-rate, false-alarm rate, true-negatives, bias, and sensitivity (Benjamin, Diaz, & Wee, 2009; Lynn &

Barrett, 2014). Accuracy was calculated as the number of correct alert classifications over total alert classifications (40). The individual differences in accuracy of the alert classifications can be viewed in Table, 3.6 along with the accuracy per alert type (i.e., one through five severity level).

Table 3.6. *Descriptive Statistics HCI*

	Hit Rate	FP Rate	CIAT Accuracy	Total Alert Response Time	Tool-Use	Alert1	Alert2	Alert3	Alert4	Alert5
Valid	16	16	16	15	15	15	15	15	15	15
Missing	0	0	0	1	1	1	1	1	1	1
Mean	0.84	0.32	0.69	19.73	145.33	0.89	0.77	0.52	0.58	0.93
Std. Deviation	0.30	0.19	0.17	9.96	61.83	0.15	0.24	0.29	0.22	0.26
Range	1.00	0.65	0.57	32.00	295.00	0.40	0.75	0.82	0.73	1.00
Minimum	0.00	0.00	0.38	5.00	20.00	0.60	0.25	0.09	0.09	0.00
Maximum	1.00	0.65	0.95	37.00	315.00	1.00	1.00	0.91	0.82	1.00

The overall mean accuracy rate for the cyber performance was 69 percent. The mean accuracy of each alert per threat-severity level from Table, 3.6 was as follows. There was high accuracy for the low severity alerts (alert security level one) at a 98 percent success level. Those alerts included failed logins and adware detected. The alerts with a severity of two had an accuracy at 78 percent. Those alerts entailed ICMP Timestamp Request, FTP Improper Port Specified, and Fragmented ICMP Packet. However, alerts with a three and four severity-level had lower accuracy. The level three severity alerts were classified at chance level at 53 percent. Those alerts included the Send Mail Recon Attack, TCP Port Sweep, ICMP Flood, TCP Null Port Scan, TCP Network Sweep, TCP High Port Scan. The category four severity alerts were classified a little over chance level at 60 percent. Those alerts included the FTP Improper Port Specified, FTP Format String Attack, Wiz attack on SMTP Server, Smurf Attack, and Ping of Death. Finally, the alerts with a severity level of five had an accuracy at 80 percent. The alert was a FTP Heap Corruption Attack. The mean RT for total alert responses was 19 minutes.

One problem with using accuracy is that it does not tell us how the analyst classified the more costly true-positive threats. Therefore, the methods of Signal Detection Theory was applied in the current study to calculate the performance measures for alert classifications (D. M. Green et al., 1966; Kornbrot, 2006). Accordingly, the 40 classified alerts were analyzed using the number of hits (i.e., number of true-positive (TP) alerts classified as true-positives), misses (i.e., number of true-positive alerts that the analyst classified as benign), false-positive alerts (i.e., number of benign alerts that the analyst classified as a threat), and correct rejections (i.e., number of benign alerts that the analyst classified as benign) for each of the 16 analysts. That data is displayed in Table, 3.7.

Table 3.7. *Descriptive Statistics for SDT*

	Hits	Misses	Correct Rejections	False Positives
Valid	16	16	16	16
Missing	0	0	0	0
Std. Deviation	0.60	0.60	7.19	7.18
Minimum	0.00	0.00	13.00	0.00
Maximum	2.00	2.00	38.00	25.00
Sum	27.00	5.00	414.00	195.00

Since all of the scores besides the hit-rate and accuracy were redundant, only hit-rate also known as Recall = TP/TP + FN (i.e., the probability of responding yes on signal trials) and accuracy were analyzed further. The false-positive rate was kept in the analysis to show its precision. In regards to the hit-rate, in total there was 32 true-positives (two true-positives per person). The total true-positives detected was 27 out of 32. There was a total of 609 false-positive threats (38 per trial) and 414 of them were correctly rejected as such. The hit-rate can be seen in Table, 3.6. The analysts correctly detected 84 percent (m = .84) of the malicious events (i.e., the hit-rate or signal) and missed 5 out of 32 (16 percent) of them. When considering all IDS experts on Purdue campus, that rate for missing true-positive alerts would be considered as poor when considering their cost. The mean false-positive rate (i.e., noise or false-alarms) was 32 percent (m = 32) for the IDS task. The false-alarm rate is higher in comparison to other cyber studies (Ben-Asher & Gonzalez, 2015; Rajivan et al., 2013), but here a more representative signal (five percent) to noise (95 percent) ratio was used, which brings the false-positive rate upward (Lynn & Barrett, 2014). The over-all error seen here for both signals (i.e., targets) and false-positives under that low of a prevalence rate for targets, is similar to what is seen in laboratory based visual search tasks for targets under similar conditions (Wolfe et al., 2005). When targets are present on just one percent of trials in that study (Wolfe et al., 2005) the observers missed 30 to 40 percent of them.

The hit-rate and false-alarm rate pairs can yield two additional measures of cyber performance to include the analysts' sensitivity and response bias, but they were not able to be computed due to normality/equality of variance being likely violated in the data due to low signals (hit-rate), and the fact that non-parametric SDT measures have been identified (N. A. Macmillan & Creelman, 1990) as having serious issues when bias is non-neutral, like it is in the current study. Therefore, analysis of those measures were not appropriate for the current data set. The overall response bias can still be seen as liberal by the high amount false-positives classified in the task.

Lastly, a specific measure of cyber performance/behavior (Motowildo et al., 1997) was taken in the current study. That was the amount of tools that were used by the analyst to investigate alerts/threats. The tool-use per analyst represented the amount of direct evidence required to make the threat decisions. The tools that were used by the analyst within the IDS software to investigate the alerts (i.e., TabSelect, Query, IPSelect, and PcapFrameSelect) was recorded and then measured by summing the amount used per trial/ per participant. The measure was taken because the work framework by Motowildo et al. (1997) suggests cognitive ability most directly affects specific work-task behaviors and habits, that then affect work-performance. Furthermore, tool-use is considered a primary behavior of an Incident Responder (D'Amico et al., 2005). Tool-use (Emmanuel et al., 2015) and information accumulation differences (Ben-Asher & Gonzalez, 2015) have also been identified in the cyber literature as contributing to cyber performance success in an IDS (Ben-Asher & Gonzalez, 2015; Emmanuel et al., 2015). The mean number of tools used to investigate all threats in the current study was (m = 145) with a range of 295 tools used, which can be seen in in Table, 3.6. That means the analyst switched between various tools on average 145 times to classify 40 alerts.

3.8 Hypothesis

The purpose of the study is to test one hypothesis: That hypothesis is:

H_0: Individual differences in attention control predicts cyber performance of network analysts.

3.9 Units and Sampling

The following section discusses the units (i.e., individuals) being tested and the sample chosen for the study. It also covers the variables being tested and evaluated, as well as what will be considered a successful test.

3.9.1 Sample

First, the current study was approved by the Purdue University Institutional Review Board, under Protocol: 1703018965. 17 participants were recruited for the study from Purdue University, over the summer of 2020. Data was not correctly saved for one participant. The 16 participants were recruited from an online advertisement that was hosted by Purdue Today that sought cyber analysts with IDS expertise on campus. Additionally, the Purdue Polytechnic Institute, the Purdue University business office, and the Center for Education and Research in Information Assurance and Security (CERIAS) business office sent out a recruitment e-mail to students and professors listed as being registered in cyber/technology courses. Finally, a recruitment e-mail was sent to the director of Information Technology at Purdue (ITaP) requesting participation in the study. The individuals interested in participating in the study e-mailed the main researcher through the provided contact information. The participant was then pre-screened for specific IDS expertise as a requirement to participate in the study. If the individual had the required IDS expertise they arranged a time to participate in the study.

The qualified participants voluntarily signed informed consent documents before beginning the study (Appendices A) and then completed a demographics questionnaire (Appendices B), followed by the IDS task, and lastly they completed the cognitive tests. The testing times were from 10am to 5pm, in the Cyber Forensics Laboratory at Purdue University. The testing was conducted by a research assistant; under the direction of the primary researcher. The testing took one to 1.5 hours to complete in one session. The participants were each given 15 dollars in exchange for their participation. The data from a total of 16 participants was retained for analysis. Although the sample size is low it is not uncommon for individual differences research in the cyber defense domain, given the noted challenge with obtaining incident responders (Dutt et al., 2013). To get around that challenge many performance studies use cognitive modeling (Dutt et al., 2013) rather than collecting data from actual analysts, and/or they commonly resort to using samples of students without cyber experience.

The sample for the current study consisted of four females (33 percent) and 12 males which happened to be representative to the cyber security work-force population where women currently account for about one quarter (24 percent) of the overall workforce (Dhamija, n.d.). The frequency for gender for the current study can be seen in Table, 3.8.

Table 3.8. *Frequencies for Gender*

Gender	Frequency	Percent
Female	4	25
Male	12	75
Missing	0	0
Total	16	100

Further, the sample had a wide age range from 18 to 64 years old, which is a far less restricted range than any cyber defense study presented in the literature review. The frequency for age is given in Table, 3.9.

Table 3.9. *Frequencies for Age*

Age	Frequency	Percent
18-25	7	43
26-33	3	18
34-43	3	18
44-64	3	18
Missing	0	0
Total	16	100

The sample included nine analysts with a graduate level education and seven analysts with a undergraduate level education, while most cyber analysts in the work-force have obtained a computer science degree (Svenmarck, 2020). The frequencies for their education are in Table, 3.10. The sample had a range of one to 10 years of specific IDS expertise, another sample range that is seldom seen in cyber defense human factors research. The IDS expertise is required due to the complexity of a real-world IDS task and for obtaining valid performance scores. The frequencies for IDS expertise in years is displayed in Table, 3.11.

Table 3.10. *Frequencies for Education*

Education	Frequency	Percent
Undergraduate	7	43
Masters	2	12
Phd	7	43
Missing	0	0
Total	16	100

Table 3.11. *Frequencies for Years of IDS Experience*

Years of IDS experience	Frequency	Percent
1	4	25
2	3	18
3	3	18
4	3	18
6	2	12
10	1	6
Missing	0	0
Total	16	100

Lastly, in regards to cyber expertise in general, the sample included seven cyber experts with six to 11 years of cyber expertise, six novice with one to two years of cyber expertise and three intermediates with three to five years of cyber expertise. The frequencies for cyber expertise is in Table, 3.12.

Table 3.12. *Frequencies for Cyber Domain Expertise*

Cyber Domain Expertise	Frequency	Percent
0-2 Novice	6	37
3-5 intermediate	3	18
6-10 Expert	7	43
Missing	0	0
Total	16	100

Overall, the sample is diverse with wide demographic ranges, that in turn increases the reliability of the measures (Cooper et al., 2017). For the current study, individual differences research is more concerned with obtaining precise/accurate scores of each individual per measure, than with increasing sample size for more statistical power to observe effects. The focus is more towards the measures because the reliability of them limits the strength of the hypothesized correlations (Cooper et al., 2017; Cronbach, 1957). For instance, the correlations between a test and other measures reduce as a function of the square root of reliability (Nunnally, 1978). Therefore, instead of increasing sample size in the current study, it applied cognitive ability measures designed for individual differences research. In those measures the amount of trials are increased in order to reduce error, leading to more reliable scores. In turn, that increased the hypothesised correlations of the current study.

Overall, the current study took great strides to obtain reliable and precise scores for both the cyber task and cognitive ability measures, and to obtain a diverse sample of cyber analysts. Therefore, the hypothesis tested on the small subset from the population reached high, statistically significant results that are reliable enough to generalize to a much larger population. The alpha levels for testing the hypothesis was set at the 0.05 level. The next section discusses the variables determined for the study.

3.9.2 Variables

There were three cognitive ability measures representing attention control that served as the predictor variables in the study; working memory, task-switching (flexibility) and visual attention via the attention networks. The criterion variable for the study included four measures of cyber performance; hit-rate, RT, accuracy and tool-use. The correlations between those measures will determine which cognitive ability contributes most to cyber performance, which then lead to their inclusion into a linear regression model.

3.9.3 Originality

Human computer interaction (HCI) research has largely used an experimental approach to study the factors that affect cyber performance, while the current study uses an individual differences approach under a work performance framework (Hunter, 1983; Motowildo et al.,

1997) to properly study the contributions of cognitive ability on cyber performance. The cognitive ability contribution to cyber performance is a neglected factor in cyber defense research, even though it is directly associated with both knowledge and work performance (Hunter, 1983). The study is original because the cognitive abilities that mostly underlie cyber performance success have not yet been scientifically identified. Since cognitive ability is the best predictor of job performance (Schmidt & Hunter, 2004), the identification of them is very beneficial and original for cyber defense work selection purposes (Cronbach, 1957). Lastly, the study is original because of its methodology. It uses actual network analysts with a wide range of expertise, age, and education levels which is hard to come by in cyber performance research. Most importantly, the analysts all had expertise using an incident detection system, which allows for more reliable and valid scores rather than resorting to the more popular method of pooling students without any experience using one. Further, the work sample for the current study is a representative primary task for the Incident Responder role, and it has been used in previous research with ample success, whereas many cyber performance studies (Emmanuel et al., 2015; Silva et al., 2014) create their own task rather than validating an existing one, and most of them are not specific to an actual cyber defense work-role (Champion et al., 2012). Finally, the study applied objective measures of both cognitive ability and task performance, while most of the cyber performance research has used subjective measures of cognitive ability to investigate its role on performance (Champion et al., 2012; Jøsok et al., 2019; Knox et al., 2017; McIntire et al., 2013; Sawyer et al., 2015). The subjective measures may serve their purpose in those studies but they do not measure the actual behavior of the cyber analyst.

3.10 Summary

The chapter provided the methods used in the research study including the reliability of the tests. The next chapter presents the results of the study.

CHAPTER 4. ANALYSIS AND RESULTS

The following chapter covers the analysis and results of the study. The human-computer interactions from each participant that performed in the CIAT simulator was collected and analyzed as measures of cyber performance, and then their cognitive ability measures were analyzed. The correlations between individual differences in cyber performance and cognitive ability will answer the hypothesis in part; by identifying the cognitive ability that most influences cyber performance. The cognitive ability with the highest correlation with cyber performance was then taken into a standard linear regression model in order to fully answer the hypothesis; that cognitive ability can predict cyber performance.

4.1 Correlational Analysis

To begin answering the hypothesis of the current study, the bi-variate Pearson correlation coefficients between the measures of cyber performance and the measured cognitive abilities is reported in the following subsections, starting with accuracy. In case of non-normality, Spearman correlations (rs) were calculated instead of Pearson correlations (r). First, normality is addressed. All variables were not normally distributed, as assessed by Shapiro-Wilk's test ($p < .05$), for performance measures seen in Table, 4.1. and for ability measures in Table, 4.2. The significantly non-normal variables were: hit-rate and working memory forwards. The Pearson and Spearman correlation results for the entire study are presented in Appendices C.

Table 4.1. *Shapiro-Wilk's Test for Performance Measures*

	Hit-Rate	FP Rate	CIAT (P)Correct	Tool-Use
Valid	16	16	16	15
Missing	0	0	0	1
Shapiro-Wilk	0.587	0.967	0.959	0.866
P-value of Shapiro-Wilk	< .001	0.794	0.648	0.030

Table 4.2. *Shapiro-Wilk's Test for Cognitive Ability Measures*

	Orient	Alert	EF	WM(F)	WM(B)	AmbiSwRTCost	SwitchRT	SwitchCoRT	Spontitchrate
Valid	16	16	16	16	16	14	15	15	16
Missing	0	0	0	0	0	2	1	1	0
Shapiro-Wilk	0.941	0.878	0.913	0.785	0.928	0.949	0.915	0.906	0.958
P-value of Shapiro-Wilk	0.365	0.037	0.128	0.002	0.223	0.546	0.163	0.117	0.624

4.1.1 Accuracy Correlations with Performance

The Pearson correlations between cognitive ability and cyber performance accuracy can be seen in Table, 4.3. Accuracy (i.e., percent of correctly classified alerts) in the cyber task negatively correlated with orienting of attention from the Attention Network Task (ANT) (r(16) = -0.654, p = .006).

Table 4.3. *Pearson's Correlations for Accuracy*

Variable		CIAT Accuracy	Orienting	Switching(RT)	Switch-Costs(RT)
1. CIAT Accuracy	n	–			
	Pearson's r	–			
	p-value	–			
2. Orienting	n	16	–		
	Pearson's r	-0.654	–		
	p-value	0.006	–		
3. Switching(RT)	n	15	15	–	
	Pearson's r	-0.620	0.218	–	
	p-value	0.014	0.436	–	
4. Switch-Costs(RT)	n	15	15	15	–
	Pearson's r	-0.599	0.463	0.842	–
	p-value	0.018	0.082	$< .001$	–

That association implied that as alert classification accuracy increases, the cost of orienting decreases. Additionally, accuracy negatively correlated with multiple conditions in the switching task (i.e., Stabflex task) (Armbruster et al., 2012) as follows. Accuracy correlated negatively with the task switching (RT) condition (r = -0.620, p = .014). Accuracy further negatively correlated with the associated switching costs (r = -0.599, p = .018). The scatter-plots for those x and y pairs can be seen in Figure, 4.1.

Accuracy did not correlate with any of the error conditions in the Stabflex task, which makes sense because the reliability of the error conditions was recently reported as low (Kraft et

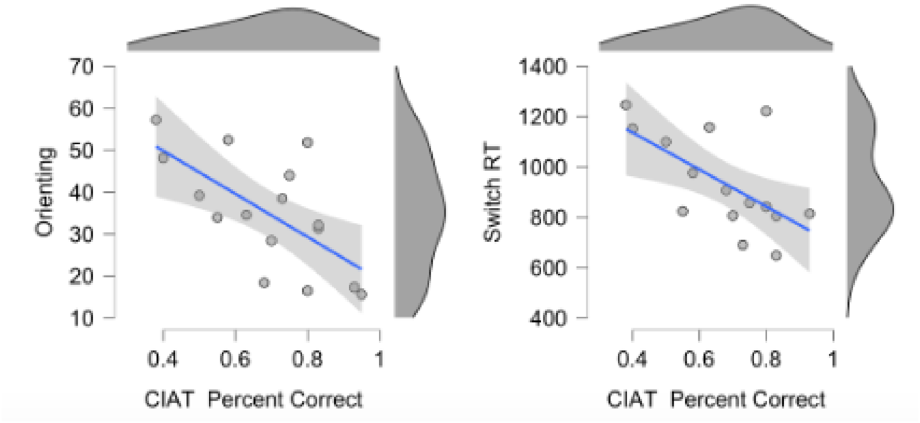


Figure 4.1. Scatter-plots Between CIAT Accuracy and Orienting and Switching(RT)

al., 2020). The correlations suggest that an individuals flexibility and spatial attention (i.e., orienting ability) contribute to cyber performance accuracy, in the predicted direction which partly answers the hypothesis of the current study. The next subsection breaks accuracy apart via methods of SDT (D. M. Green et al., 1966; Stanislaw & Todorov, 1999) in order to measure the hit-rate separately from the false-positive rate in the IDS task.

4.1.2 Correlations Between SDT Measures of Cyber Performance and Cognitive Ability

The bi-variate Spearman correlations (rs) were calculated between the hit-rate measure of cyber performance and the measured cognitive abilities. First the hit-rate for true-positive alert classifications in the cyber task negatively correlated with the alerting network (rs = -0.524, p = .037, n = 14) from the ANT measure, as seen in Table 4.4. The correlation indicates that more efficient vigilance is associated with the ability to detect a signal or true-positive threat in the IDS task. After visually verifying the associations in the scatter-plot in Figure, 4.2 more data is required to confirm the relationship. The cognitive ability contribution to the false-positive rate is discussed next.

The Pearson correlations between cognitive ability and the false-positive rate can be seen in Table, 4.5 for comparing to accuracy in Table, 4.3. The false-positive rate correlated in an opposite direction (positive) than did accuracy for the same cognitive variables, but more significantly. The false-positives correlated positively with the orienting network (r = 0.645, p =

Table 4.4. *Correlation Table for Hit-Rate*

Variable		Hit Rate	Alerting
1. Hit Rate	n	–	
	Pearson's r	–	
	p-value	–	
	Spearman's rho	–	
	p-value	–	
2. Alerting	n	16	–
	Pearson's r	-0.73	–
	p-value	1.20e-3	–
	Spearman's rho	-0.52	–
	p-value	0.04	–

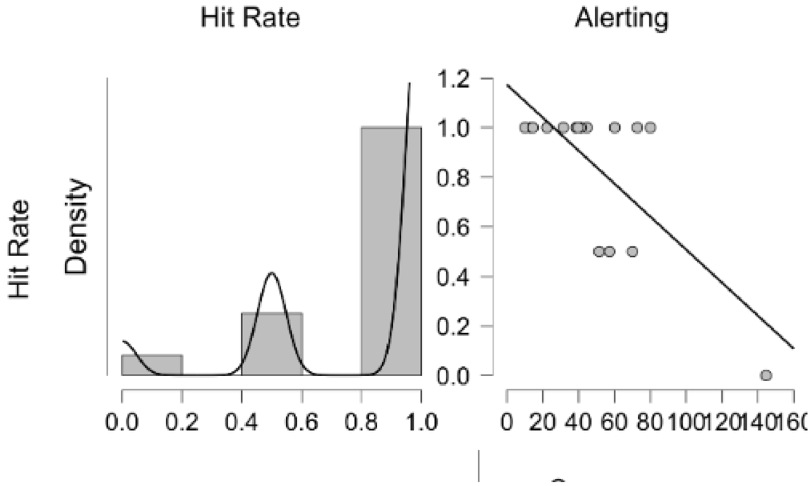


Figure 4.2. A Scatter-plot for the Hit-Rate and Alerting Correlation

.007, n = 15) from the ANT measure (Fan et al., 2002). It also positively correlated with switching costs (RT) (r = 0.602, p = .018) from the task-switching measure (Armbruster et al., 2012). Those correlations indicate that increasing classification errors (false-alarms) in the IDS is associated with higher cognitive costs. Finally, the false-positive rate located in Table, 4.5 negatively correlated with frequency of tools used to investigate alerts (r = -0.544, p = .036). That finding indicates that more tools used to investigate alerts, is associated with lower cyber performance error (i.e., false-positive rate).

Table 4.5. *Pearson's Correlations for False-Positive Rate*

Variable		Orienting	Switching(RT)	Switch-Costs(RT)	FP Rate
1. Orienting	n	–			
	Pearson's r	–			
	p-value	–			
2. Switching(RT)	n	15	–		
	Pearson's r	0.218	–		
	p-value	0.436	–		
3. Switch Costs(RT)	n	15	15	–	
	Pearson's r	0.463	0.842	–	
	p-value	0.082	$< .001$	–	
4. FP Rate	n	16	15	15	–
	Pearson's r	0.645	0.626	0.602	–
	p-value	0.007	0.013	0.018	–

4.1.3 Correlations Between Specific Behaviors of Cyber Performance and Cognitive Ability

The frequency of tools used to make a decision (a criterion) when investigating alerts is a specific behavior of task performance that correlated with the orienting network of the ANT ($r(15) = -0.691$, $p = .004$) to suggest that quicker orienting ability may lead to more tool-use. The frequency of tools used is also associated with increased alert response time, which is represented by the positive correlation coefficient between them ($r(15) = 0.616$, $p = .014$). Tool-use also associated with cyber performance accuracy ($r(15) = 0.540$, $p = 0.03$) indicating that accuracy increases with more tool-use. The scatter-plots of those variables are displayed in Figure, 4.3 in order to visually verify the associations. The Pearson correlations between those variables can be seen in Table, 4.6.

The associations conclude that individual differences in spatial attention though the attention networks (Fan et al., 2002) underlies the search for evidence behavior in an IDS. The finding suggests that more efficient spatial attention may allow for more tool-use when investigating alerts.

Table 4.6. *Pearson's Correlations for Tool-Use*

Variable		Tool-Use	CIAT Accuracy	Orienting	Alert Response Time
1. Tool-Use	n	–			
	Pearson's r	–			
	p-value	–			
2. CIAT Accuracy	n	15	–		
	Pearson's r	0.540	–		
	p-value	0.038	–		
3. Orienting	n	15	16	–	
	Pearson's r	-0.691	-0.654	–	
	p-value	0.004	0.006	–	
4. Alert Response Time	n	15	15	15	–
	Pearson's r	0.616	0.100	-0.451	–
	p-value	0.014	0.723	0.092	–

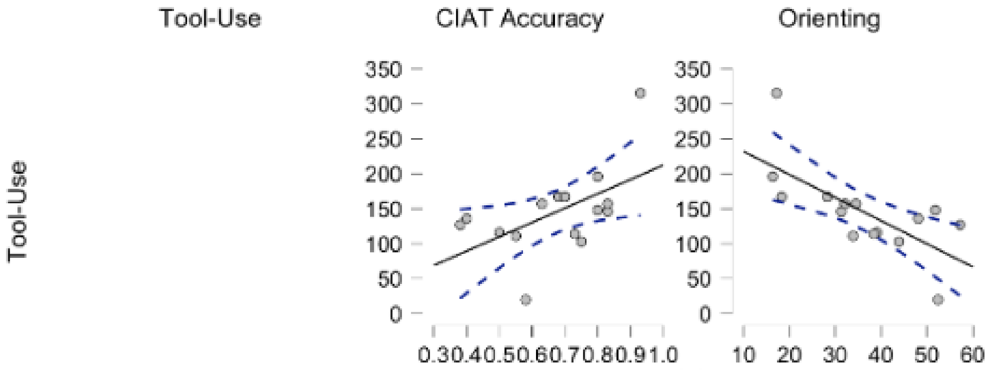


Figure 4.3. Scatter-plots For Tool-Use on CIAT Accuracy and Orienting

4.1.4 Correlation Analysis Summary

The findings presented thus far support the hypothesis of the current study, under the individual differences in job performance framework (Motowildo et al., 1997) that suggests cognitive ability directly impacts specific behaviors and habits; that then impacts performance. That relationship can be seen in the findings of the current study between spatial attention ability, tool-use behaviors and cyber performance accuracy.

4.2 Linear Regression

The highly significant cognitive ability in association with cyber performance was taken into a linear regression model in order to fully answer the hypothesis; individual differences in cognitive ability predict cyber performance success. A linear regression model was used to explore how much the predictor variables (i.e., orienting and task-switching) can predict/explain cyber performance accuracy (a continuous variable) for 16 cyber analysts. The descriptive statistics for the variables can be seen in Table, 4.7.

Table 4.7. *Descriptive Statistics*

	CIAT Accuracy	Orienting	SwitchingRT
Valid	16	16	15
Missing	0	0	1
Mean	0.69	34.90	935.65
Std. Error of Mean	0.04	3.38	49.96
Std. Deviation	0.17	13.53	193.51
Minimum	0.38	15.50	646.79
Maximum	0.95	57.20	1245.76

First, the outliers that indicated experimental error were removed, but not the outliers that were clearly due to variability in the measurements since linear regressions are sensitive to outliers. That was done through viewing histograms and scatter-plots of the variables. Further, Table 4.8 shows that the standardized residuals do not exceed -3.00 to 3.00, so no outliers were present. The Bi-variate Pearson correlations between task-switching (RT), orienting and cyber performance computed for hypothesis one can be seen again in Table, 4.9. The size of the effect represented by each correlation was determined using Cohen's (1988) criteria in which (r = 0.10) represents a small effect, (r = 0.30) represents a medium effect and (r = 0.50) represents a large effect. Table, 4.9 shows that cyber performance demonstrated a strong bi-variate correlation with task switching (r = -0.620, p = .007; a large effect), along with orienting (r = -0.654, p = .006; a large effect). The signs for each of the Pearson correlations was in the predicted direction, and each correlation was statistically significant with alpha set at (p = .05). Participants with more efficient (e.g., attention (i.e., orienting and task-switching), tended to have higher cyber performance scores.

Table 4.8. *Residuals Statistics*

	Minimum	Maximum	Mean	SD	N
Predicted Value	0.411	0.841	0.673	0.127	15
Residual	-0.186	0.180	4.860e-18	0.103	15
Std. Predicted Value	-2.065	1.325	-2.808e-16	1.000	15
Std. Residual	-1.752	1.830	0.027	1.047	15

Table 4.9. *Pearson's Correlations*

Variable		CIAT Percent Correct	Switch(RT)	Orienting
1. CIAT Percent Correct	n	–		
	Pearson's r	–		
	p-value	–		
2. Switch(RT)	n	15	–	
	Pearson's r	-0.620	–	
	p-value	0.014	–	
3. Orienting	n	16	15	–
	Pearson's r	-0.654	0.218	–
	p-value	0.006	0.435	–

The multiple regression analysis in Tables 4.10, and in 4.11, and in 4.12, showed that there was a statistically significant relationship between (a) cyber performance scores (CIAT percent correct/accuracy) and (b) the 2 predictor variables (i.e., orienting and task switching) taken as a set, R2 = 0.60 percent, adjusted R2 = 0.53 percent, F(2, 12) = 9.139, $p < .0005$.

Table 4.10. *ANOVA*

Model		Sum of Squares	df	Mean Square	F	p
H_1	Regression	0.225	2	0.113	9.139	0.004
	Residual	0.148	12	0.012		
	Total	0.373	14			

Table 4.11. *Model Summary - CIAT Percent Correct*

										Durbin-Watson		
Model	R	R^2	Adjusted R^2	RMSE	R^2 Change	F Change	df1	df2	p	Autocorrelation	Statistic	p
H_0	0.000	0.000	0.000	0.163	0.000		0	14		0.299	1.334	0.174
H_1	0.777	0.604	0.538	0.111	0.604	9.139	2	12	0.004	0.294	1.373	0.144

Table 4.12. *Coefficients*

Model		Unstandardized	Standard Error	Standardized	t	p
H_0	(Intercept)	0.673	0.042		15.950	$< .001$
H_1	(Intercept)	1.299	0.156		8.345	$< .001$
	Orienting	-0.006	0.002	-0.479	-2.573	0.024
	Switching RT	-4.356e-4	1.572e-4	-0.516	-2.771	0.017

The strength of the relationship meets the criteria of a large effect according to Cohen (1988), indicating that orienting and task switching accounted for 53 to 60 percent of the variance in cyber performance scores using the model. The sign of the multiple regression coefficient for each predictor variable was in the predicted direction. For instance, as orienting decreased, cyber performance increased; a negative association. The predictor variables each displayed a multiple regression coefficient that was statistically significant, as follows: Orienting (b = -.6, p = .024), and Set-shifting (b = -.04, p = .017). That means that as orienting (X) increases 1 millisecond, cyber performance Y decreases .6 percent. As switching RT increases 1 millisecond, cyber performance decreases by .04 percent. Those variables can be entered into the linear equation model to predict cyber performance scores. The partial regression plots between CIAT accuracy and orienting, can be seen in Figure, 4.4 and the partial regression plot between CIAT accuracy and task-switching can be seen in Figure, 4.5. The rest of the linear regression assumptions are now addressed.

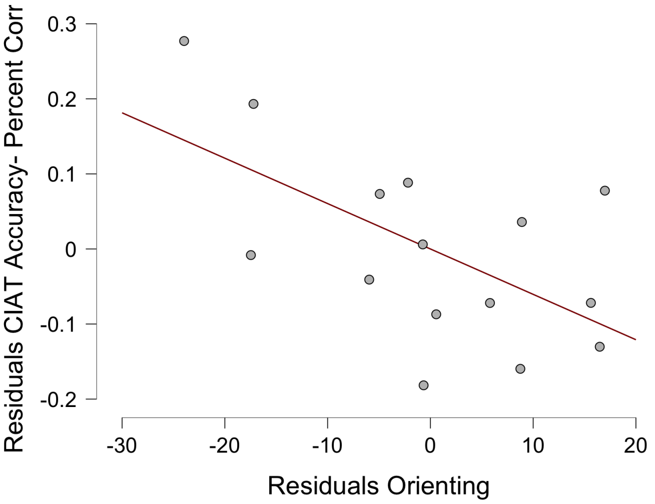


Figure 4.4. Scatter-plots for CIAT Accuracy on Orienting and Task-Switching(RT)

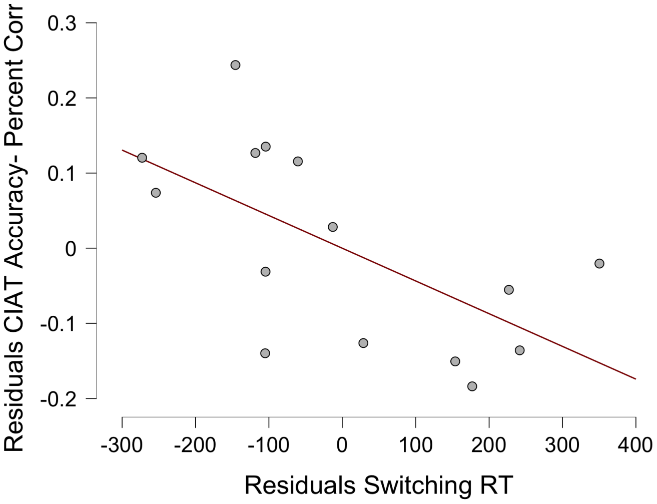


Figure 4.5. Scatter-plots for CIAT Percent Correct on Switching(RT) and Task-Switching(RT)

First, the Durbin Watson Statistic seen in the model summary of Table, 4.10 checks for correlations between residuals (i.e., error). That result was within an acceptable range above 1 and below 3 (Durbin Watson = 1.373), therefore the assumption of no-autocorrelation of residuals has been met. There was homoscedasticity, as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values. The variance inflation factor (VIF) value was a low 1 and none of the tolerance values were above 10, therefore the assumption of no multicollinearity has been met. The quantile-quantile plot (Q-Q plot) in 4.6 shows that the standardized residuals fit nicely along the diagonal, suggesting that the assumptions of normality have also not been violated.

At this point, all of the assumptions of the regression model were met for orienting, set-shifting (RT) and cyber performance (i.e., accuracy).

4.3 Analysis Conclusions

The primary goal of the current study was to investigate the cognitive abilities contribution to cyber performance using a sample of experienced cyber analysts with wide variance in expertise. The current study hypothesised that cognitive ability can predict cyber performance success in an incident detection system (IDS). The hypothesis was supported as true

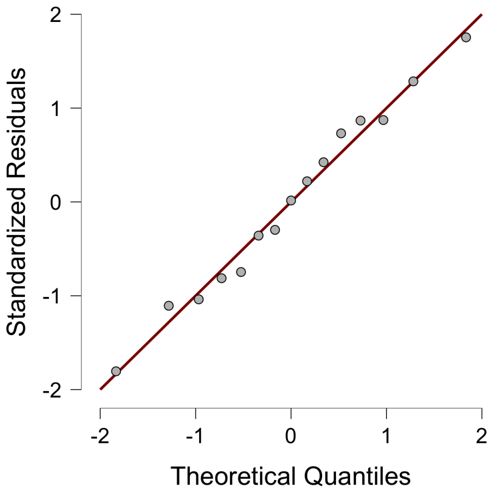


Figure 4.6. QQ Plot

in the highly significant negative correlations seen between the spatial attention networks and task-switching ability (i.e., flexibility) measures on IDS performance. In further support of the hypothesis, the analysts' evidence accumulation (i.e., a specific behavior) (Motowildo et al., 1997) through tool-use in the IDS was strongly associated with orienting ability in the predicted direction. A higher orienting ability may have allowed some analysts to obtain a higher level of evidence per alert (i.e., evidence accumulation) through their use of more tools prior to making the threat decisions. More tool-use was also correlated with an overall higher cyber performance; a higher accuracy. The correlations indicate that more efficiency in the identified cognitive processes mostly influence the differences in cyber performance.

The hypothesis is further supported as true in the findings from a linear regression model which demonstrated that orienting of attention and flexibility predict 53 percent to 60 percent of the variance in cyber performance scores in an incident detection task. The cognitive ability contribution to performance is what is similarly seen between General Mental Ability (GMA) and work-tasks ($r = 0.51$) in selection research (Hunter, 1983; Schmidt & Hunter, 2004; Schmidt et al., 1986), and for more complex work-tasks (Schmidt & Hunter, 2004). On the other hand, many of the cognitive ability measures did not associate with cyber performance. First, the error conditions from the task-switching measures (i.e., Stabflex) did not influence cyber performance, for one reason being that its reliability was recently discovered as low (Kraft et al., 2020). Second, the working memory task (Woods et al., 2011) did not associate with performance measures however, working memory in the Stabflex measure (Armbruster et al., 2012) did. The correlations between cyber performance and flexibility means that in order to perform well, the analyst must be able to maintain information in working memory as well as task-switch with flexibility through the cyber environment (Miyake & Friedman, 2012). The WM construct (Baddeley & Hitch, 1974) and task (i.e., digit span) by itself may not be difficult enough to produce individual differences in a group of cyber analysts.

Although not all ability measures were successful in regards to its hypothesised association with cyber performance, the objective of the study to identify underlying cognitive ability contribution to more successful cyber performance was achieved in the individual differences study and the hypothesis was supported. The next section of the dissertation discusses the conclusions and implications of the research and what it means to cyber defense.

CHAPTER 5. CONCLUSIONS

The purpose of the research was to answer the question; how do individual differences in cognitive ability contribute to the cyber performance of network security analysts? To investigate the research question, one hypothesis was tested. The first part of the hypothesis was tested by investigating which cognitive ability contributes most to cyber performance. For that, the correlations between attention control and cyber performance scores was investigated. The results of the correlational analysis showed that individual differences in flexibility and spatial attention contribute most to the differences in IDS work performance. The second part of the hypothesis was tested by investigating if individual differences in attention control can predict the cyber performance of network analysts. The results of the linear regression model demonstrated that orienting of attention and flexibility predict 53 to 60 percent of the variance in cyber performance scores. Those findings have important implications for cyber defense.

The major finding from the hypothesis; that spatial attention and flexibility contribute most to cyber performance confirms the cyber defense research (Greenlee et al., 2016; Jøsok et al., 2019; Knox et al., 2017) that largely assumes those abilities associate most with cyber performance; and it also contributes to the experimental research that found vigilance to strongly effect cyber task performance (McIntire et al., 2013; Sawyer et al., 2015), which strengthens the findings of the current study. Further, the findings confirm cyber frameworks (Andrade et al., 2018) that posit executive function and flexibility as a highly significant contributor to cyber performance success without scientific evidence. The findings also contribute to the human factor models (Endsley, 1995; Jøsok et al., 2019; Knox et al., 2017) that pinpoint various cognitive abilities that underlie cyber performance, but have not scientifically identified a single one that majorly contributes to cyber performance success. Lastly, it contributes to the human cognitive decision models (Dutt et al., 2013; Funke et al., 2016) that include working memory ability and various other cognitive abilities to simulate the behavior of the analyst without knowing the main contributors to cyber performance.

A practical application of the correlational findings would be to include the identified cognitive ability in the interface design research for cyber defense systems. Individual differences in flexibility and spatial attention behavior represent important aspects of analyst performance and should be considered in interface design of cyber defense software to improve performance. For instance, software can be tested with the identified abilities from the current study with the aim to alleviate the cognitive load on those processes so that all analysts can perform well. That would require using treatments (e.g., different interface designs) and then allocating individuals to the different treatments depending on their cognitive ability level in association with performance (Cronbach, 1957). Cronbach (1957) argued that, ”ultimately we should design treatments, not to fit the average person, but to fit groups of students with particular aptitude patterns. Conversely, we should seek out the aptitudes which correspond to (interact with) modifiable aspects of the treatment” (Cronbach, 1957, p.681). The current study identifies the cognitive ability that interacts most with cyber performance for further performance research efforts. In particular, the tool-use requirement appears to be an important aspect of cyber performance where a software aid would be beneficial because of its seemingly high cognitive demand.

Additionally, the current study found that higher tool-use (a specific behavior in the IDS) correlated with lower spatial attention costs (i.e., a quicker orienting ability), and that it also associated with higher cyber performance. First, that finding builds on the cyber performance research (Ben-Asher & Gonzalez, 2015; Emmanuel et al., 2015) that compared groups of analysts and found separately that experts use more network tools (Emmanuel et al., 2015) and accumulate more evidence than the novice prior to making attack decisions (Ben-Asher & Gonzalez, 2015), by identifying the cognitive ability that contributes to both information search and evidence accumulation (D’Amico et al., 2005); the actual primary task of the Incident Responder. The finding also supports research from applied visual attention research (Wolfe et al., 2005) where the extent of the search for evidence (i.e., targets) likely depends on the individual (Wolfe et al., 2005); identified here in the cyber domain as individual differences in visual attention control. Lastly, the finding supports the work performance framework (Hunter, 1983; Motowildo et al., 1997) which posits that individual differences in cognitive ability directly impacts task behavior and task-habits; that contribute to job performance success (Motowildo et al., 1997).

The other major finding from the hypothesis; that orienting of attention and flexibility predict 53 to 60 percent of the variance in cyber performance scores has important practical implications for the cyber defense work force. First, the findings from the hypothesis scientifically identified the cognitive ability contribution to more successful cyber defending for the Incident Responder. Thus, it identified the missing abilities from the NICE framework (Newhouse et al., 2017; Petersen et al., 2020) that are most desirable for performing in the Incident Responder role. The NICE framework should be more inclusive of cognitive ability after considering the empirical research behind the long established work performance frameworks (Hunter, 1983; Motowildo et al., 1997; Schmidt & Hunter, 2004, 1998), and because the best work selection predictors include cognitive ability measures such as the GMA (Schmidt & Hunter, 2004), especially for highly cognitive work-roles like those in cyber defense. Similarly, but more specifically, attention control was able to predict cyber performance success in the current study. The identified cognitive abilities can now be applied for selecting talent for a more efficient cyber workforce (Schmidt & Hunter, 2004). The companies that use valid measures for hiring purposes that actually predict job performance, like those identified in the current study are drastically more efficient over time (Schmidt & Hunter, 2004). Schmidt and Hunter (1998) research makes three points regarding the benefits of using valid selection methods, “(a) the economic value of gains from improved hiring methods are typically quite large, (b) these gains are directly proportional to the size of the increase in validity when moving from the old to the new selection methods, and (c) no other characteristic of a personnel measure is as important as predictive validity” (Schmidt & Hunter, 1998, p.262).

Lastly, the observed relationships between the cognitive measures (e.g., flexibility, orienting) and the tool-use behavior in the cyber task confirms that the laboratory measures are valid, in that they are measuring the underlying cognitive processes that they are intended to measure. Therefore, the findings support the predictive validity of the ANT (Fan et al., 2002) and task-switching measure (Armbruster et al., 2012) by demonstrating that an individuals score from those measures can be used to predict cyber performance with reasonable accuracy. Overall, the results of the study indicate that the hypothesis from the current individual differences study is correct. The limitations of the study are now discussed.

The first set of limitations to the study in regards to the predictor measures was that some of them did not associate significantly with performance, but it was clear as to why. Either the measure did not produce a wide enough variance and/or the measure was not reliable enough for individual differences research. For instance, the digit-span could have been analyzed differently in order to obtain more variance for individual differences research using additional methods from Woods et al. (2011), which could have increased its reliability. The limitations pertaining to the criterion are now discussed.

A limitation regarding the cyber performance measures was in the somewhat failed application of Signal Detection (SDT) methods (Stanislaw & Todorov, 1999). Although it was seen through SDT methods that vigilance may play an important role in making true-positive threat decisions, further measures could not be obtained such as the response bias and sensitivity. The current study did find that SDT methods measured cyber performance more precisely and lead to higher correlations with cognitive ability than did accuracy. The true-positive alerts should be investigated and weighted separately from false-negative threat classifications in an IDS because of the enormous losses and/or costs for missing true-positive threats in a cyber domain. Furthermore, the response bias should also be accounted for when measuring IDS performance again because of the associated costs. There was not enough true-positive alerts in the IDS to measure bias in the current study. Another limitation pertaining to the criterion is that cognitive engagement alerts could have been included in the IDS task to ensure that the participants were participating as they should.

The last notable limitation to the current study pertains to the low sample size, but the sample size is not unusual for research studying human factors in cyber defense. The sample size could have been increased, but that would have required funding and additional IRB approval to seek analysts outside of Purdue University. A replication of the current study with more participants is suggested as a potential future study.

A future study can administer a more lengthy domain knowledge and experience questionnaire (Ben-Asher & Gonzalez, 2015) prior to performing in the cyber task, and then the associations between cognitive ability, knowledge and work/task performance (Hunter, 1983) can be investigated further. Measuring both cognitive ability and knowledge is important because knowledge or expertise alone does not guarantee performance success (Schmidt et al., 1986). Overall, the current study is a start to scientifically exploring the cognitive abilities that transfer to various aspects of cyber defense.

REFERENCES

Allport, A., & Wylie, G. (2000). Task switching, stimulus-response bindings, and negative priming. *Control of cognitive processes: Attention and performance XVIII*, 35–70.

Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks.

Andrade, R. O., Cazares, M. F., Tello-Oquendo, L., Fuertes, W., Samaniego, N., Cadena, S., & Tapia, F. (2018). From cognitive skills to automated cybersecurity.

Armbruster, D. J., Ueltzh, K., Basten, U., & Fiebach, C. J. (2012). Prefrontal cortical mechanisms underlying individual differences in cognitive flexibility and stability. *Journal of cognitive neuroscience*, *24*(12), 2385–2399.

Armbruster-Genç, D. J., Ueltzhöffer, K., & Fiebach, C. J. (2016). Brain signal variability differentially affects cognitive flexibility and cognitive stability. *Journal of Neuroscience*, *36*(14), 3978–3987.

Baddeley, A. D., & Hitch, G. (1974). Working memory. , *8*, 47–89.

Banfield, J. F., Wyland, C. L., Macrae, C. N., Münte, T. F., & Heatherton, T. F. (2004). The cognitive neuroscience of self-regulation.

Ben-Asher, N., & Gonzalez, C. (2015). Effects of cyber security knowledge on attack detection. *Computers in Human Behavior*, *48*, 51–61.

Ben-Asher, N., Oltramari, A., Erbacher, R. F., & Gonzalez, C. (2015). Ontology-based adaptive systems of cyber defense. In *Stids* (pp. 34–41).

Benjamin, A. S., Diaz, M., & Wee, S. (2009). Signal detection with criterion noise: applications to recognition memory. *Psychological review*, *116*(1), 84.

Borneman, M. M. (2018). Estimating defensive cyber operator decision confidence.

Broadbent, D. E., Cooper, P. F., FitzGerald, P., & Parkes, K. R. (1982). The cognitive failures questionnaire (cfq) and its correlates. *British journal of clinical psychology*, *21*(1), 1–16.

Campbell, S. (2016). *Evaluating the attention network test and its ability to detect cognitive decline*. Unpublished doctoral dissertation.

Campbell, S. G., O'Rourke, P., & Bunting, M. F. (2015). Identifying dimensions of cyber aptitude the design of the cyber aptitude and talent assessment. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 59, pp. 721–725).

Causse, M., Dehais, F., & Pastor, J. (2011). Executive functions and pilot characteristics predict flight simulator performance in general aviation pilots. *The International Journal of Aviation Psychology*, *21*(3), 217–234.

Champion, M. A., Rajivan, P., Cooke, N. J., & Jariwala, S. (2012). Team-based cyber defense analysis. In *Cognitive methods in situation awareness and decision support (cogsima), 2012 ieee international multi-disciplinary conference* (pp. 218–221).

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual review of psychology*, *62*, 73–101.

Cisco. (2014). *Cisco 2014 annual security report* (Tech. Rep.). www.cisco.com: Author.

Cisco. (2020). *Cisco benchmark study: Securing what's now, and what's next. 20 cybersecurity considerations for 2020* (Tech. Rep.). www.cisco.com: Author.

Cohen, M. A., Cavanagh, P., Chun, M. M., & Nakayama, K. (2012). The attentional requirements of consciousness. *Trends in cognitive sciences*, *16*(8), 411–417.

Colzato, L. S., Van Leeuwen, P. J., Van Den Wildenberg, W., & Hommel, B. (2010). Doom'd to switch: superior cognitive flexibility in players of first person shooter games. *Frontiers in psychology*, *1*, 8.

Compton, S., & Hornat, C. (2007). 802.11 denial of service attacks and mitigation. *SANS Institute InfoSec Reading Room*, 14–18.

Cooper, S. R., Gonthier, C., Barch, D. M., & Braver, T. S. (2017). The role of psychometrics in individual differences research in cognition: A case study of the ax-cpt. *Frontiers in psychology*, *8*, 1482.

Cranford, E. A., Gonzalez, C., Aggarwal, P., Cooney, S., Tambe, M., & Lebiere, C. (2020). Toward personalized deceptive signaling for cyber defense using cognitive models. *Topics in Cognitive Science*, *12*(3), 992–1011.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American psychologist*, *12*(11), 671.

D'Amico, A., Whitley, K., Tesone, D., O'Brien, B., & Roth, E. (2005). Achieving cyber defense situational awareness: A cognitive task analysis of information assurance analysts. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 49, pp. 229–233).

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Memory and Language*, *19*(4), 450.

Das, R., Karabade, A., & Tuna, G. (2015). Common network attack types and defense mechanisms. In *2015 23nd signal processing and communications applications conference (siu)* (pp. 2658–2661).

Dawson, J., & Thomson, R. (2018). The future cybersecurity workforce: Going beyond technical skills for successful cyber performance. *Frontiers in psychology*, *9*, 744.

De Jong, R. (2000). 15 an intention-activation account of residual switch costs. *Control of cognitive processes*, 357.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual review of neuroscience*, *18*(1), 193–222.

Dhamija, A. (n.d.). Favorable factors to improve gender diversity in cybersecurity industry.

Dutt, V., Ahn, Y.-S., & Gonzalez, C. (2013). Cyber situation awareness modeling detection of cyber attacks with instance-based learning theory. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *55*(3), 605–618.

D'Amico, A., & Whitley, K. (2008). The real work of computer network defense analysts. In *Vizsec 2007* (pp. 19–37). Springer.

Emmanuel, G. R., McClain, J. T., Matzen, L. E., & Forsythe, J. C. (2015). *Measuring expert and novice performance within computer security incident response teams.* (Tech. Rep.). Sandia National Lab.(SNL-NM), Albuquerque, NM (United States).

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*(1), 32–64.

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. *Psychology of learning and motivation*, *44*, 145–200.

Enns, J. T., & Richards, J. C. (1997). Visual attentional orienting in developing hockey players. *Journal of Experimental Child Psychology*, *64*(2), 255–275.

Eriksen, C. W. (1995). The flankers task and response competition: A useful tool for investigating a variety of cognitive problems. *Visual Cognition*, *2*(2-3), 101–118.

Evangelopoulou, M., & Johnson, C. W. (2015). Empirical framework for situation awareness measurement techniques in network defense. In *Cyber situational awareness, data analytics and assessment (cybersa), 2015 international conference on* (pp. 1–4).

Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., & Posner, M. I. (2005). The activation of attentional networks. *Neuroimage*, *26*(2), 471–479.

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of cognitive neuroscience*, *14*(3), 340–347.

Fechner, G. T., Howes, D. H., & Boring, E. G. (1966). *Elements of psychophysics* (Vol. 1). Holt, Rinehart and Winston New York.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human perception and performance*, *18*(4), 1030.

Fougnie, D., & Marois, R. (2007). Executive working memory load induces inattentional blindness. *Psychonomic bulletin & review*, *14*(1), 142–147.

Frost, R. (1915). A servant to servants. *North of Boston*, *24*.

Funke, G., Dye, G., Borghetti, B., Mancuso, V., Greenlee, E., Miller, B., … Vieane, A. (2016). Development and validation of the air force cyber intruder alert testbed (ciat). In *Advances in human factors in cybersecurity* (pp. 363–376). Springer.

Gillespie, C. R., & Eysenck, M. W. (1980). Effects of introversion-extraversion on continuous recognition memory. *Bulletin of the Psychonomic Society*, *15*(4), 233–235.

Good, D., & Yeganeh, B. (2012). Cognitive agility: adapting to real-time decision making at work. *OD Practitioner*, *44*(2), 13–17.

Goodall, J. R. (2011). An evaluation of visual and textual network analysis tools. *Information Visualization*, *10*(2), 145–157.

Goodall, J. R., Lutters, W. G., & Komlodi, A. (2009a). Developing expertise for network intrusion detection. *Information Technology & People*, *22*(2), 92–108.

Goodall, J. R., Lutters, W. G., & Komlodi, A. (2009b). Supporting intrusion detection work practice. *Journal of Information System Security*, *5*(2), 42–73.

Green, C. S., & Bavelier, D. (2006). Effect of action video games on the spatial distribution of visuospatial attention. *Journal of experimental psychology: Human perception and performance*, *32*(6), 1465.

Green, D. M., Swets, J. A., et al. (1966). *Signal detection theory and psychophysics* (Vol. 1). Wiley New York.

Greenlee, E. T., Funke, G. J., Warm, J. S., Sawyer, B. D., Finomore, V. S., Mancuso, V. F., … Matthews, G. (2016). Stress and workload profiles of network analysis: Not all tasks are created equal. In *Advances in human factors in cybersecurity* (pp. 153–166). Springer.

Guastello, S. J., Shircel, A., Malon, M., & Timm, P. (2015). Individual differences in the experience of cognitive workload. *Theoretical Issues in Ergonomics Science*, *16*(1), 20–52.

Gutzwiller, R. (2007). *Switch choice in applied multi-task management*. Unpublished doctoral dissertation, Colorado State University. Libraries.

Gutzwiller, R. S., Hunt, S. M., & Lange, D. S. (2016). A task analysis toward characterizing cyber-cognitive situation awareness (ccsa) in cyber defense analysts. In *2016 ieee international multi-disciplinary conference on cognitive methods in situation awareness and decision support (cogsima)* (pp. 14–20).

Harvey Jr, L. O., Hammond, K. R., Lusk, C. M., & Mross, E. F. (1992). The application of signal detection theory to weather forecasting behavior. *Monthly Weather Review*, *120*(5), 863–883.

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*(3), 1166–1186.

Hitchcock, E. M., Warm, J. S., Matthews, G., Dember, W. N., Shear, P. K., Tripp, L. D., … Parasuraman, R. (2003). Automation cueing modulates cerebral blood flow and vigilance in a simulated air traffic control task. *Theoretical Issues in Ergonomics Science*, *4*(1-2), 89–112.

Holleman, G. A., Hooge, I. T., Kemner, C., & Hessels, R. S. (2020). The ‘real-world approach’and its problems: A critique of the term ecological validity. *Frontiers in Psychology*, *11*, 721.

Huh, T. J., Kramer, J. H., Gazzaley, A., & Delis, D. C. (2006). Response bias and aging on a recognition memory task. *Journal of the International Neuropsychological Society*, *12*(1), 1–7.

Hunter, J. E. (1983). A causal analysis of cognitive ability, job knowledge, job performance, and supervisor ratings. *Performance measurement and theory*, *257*, 266.

Huq, N. (2015). Follow the data: Dissecting data breaches and debunking myths. *TrendMicro Research Paper (Sept. 2015)*.

Ishigami, Y., & Klein, R. M. (2010). Repeated measurement of the components of attention using two versions of the attention network test (ant): stability, isolability, robustness, and reliability. *Journal of neuroscience methods*, *190*(1), 117–128.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature reviews neuroscience*, *2*(3), 194–203.

Jersild, A. T. (1927). Mental set and shift. *Archives of psychology*.

Jones, K. S., Namin, A. S., & Armstrong, M. E. (2018). The core cyber-defense knowledge, skills, and abilities that cybersecurity students should learn in school: Results from interviews with cybersecurity professionals. *ACM Transactions on Computing Education (TOCE)*, *18*(3), 1–12.

Jøsok, Ø., Lugo, R., Knox, B. J., Sütterlin, S., & Helkala, K. (2019). Self-regulation and cognitive agility in cyber operations. *Frontiers in psychology*, *10*.

Kahneman, D., Treisman, A., et al. (1984). Changing views of attention and automaticity. *Varieties of attention*, *1*, 29–61.

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.

Kantner, J., & Lindsay, D. S. (2012). Response bias in recognition memory as a cognitive trait. *Memory & cognition*, *40*(8), 1163–1177.

Kantner, J., & Lindsay, D. S. (2014). Cross-situational consistency in recognition memory response bias. *Psychonomic Bulletin & Review*, *21*(5), 1272–1280.

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological bulletin*, *144*(11), 1147.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A., & Koch, I. (2010). Control and interference in task switching: a review. *Psychological bulletin*, *136*(5), 849.

Klauer, K. C., & Kellen, D. (2010). Toward a complete decision model of item and source recognition: A discrete-state approach. *Psychonomic Bulletin & Review*, *17*(4), 465–478.

Knox, B. J., Lugo, R. G., Jøsok, Ø., Helkala, K., & Sütterlin, S. (2017). Towards a cognitive agility index: the role of metacognition in human computer interaction. In *International conference on human-computer interaction* (pp. 330–338).

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in cognitive sciences*, *11*(1), 16–22.

Kokkonen, T., & Puuska, S. (2018). Blue team communication and reporting for enhancing situational awareness from white team perspective in cyber security exercises. In *Internet of things, smart spaces, and next generation networks and systems* (pp. 277–288). Springer.

Kornbrot, D. (2006). Signal detection theory, the approach of choice: Model-based and distribution-free measures and evaluation. *Perception & Psychophysics*.

Kortschot, S. W., Sovilj, D., Jamieson, G. A., Sanner, S., Carrasco, C., & Soh, H. (2018). Measuring and mitigating the costs of attentional switches in active network monitoring for cybersecurity. *Human factors*, *60*(7), 962–977.

Kraft, D., Rademacher, L., Eckart, C., & Fiebach, C. J. (2020). Cognitive, affective, and feedback-based flexibility–disentangling shared and different aspects of three facets of psychological flexibility. *Journal of cognition*, *3*(1).

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human perception and performance*, *21*(3), 451.

Lecerf, T., & Roulin, J.-L. (2009). Individual differences in visuospatial working memory capacity and distractor inhibition. *Swiss Journal of Psychology*, *68*(2), 67–78.

Lerman, D. C., Tetreault, A., Hovanetz, A., Bellaci, E., Miller, J., Karp, H., … others (2010). Applying signal-detection theory to the study of observer accuracy and bias in behavioral assessment. *Journal of applied behavior analysis*, *43*(2), 195–213.

Lynn, S. K., & Barrett, L. F. (2014). Utilizing signal detection theory. *Psychological science*, *25*(9), 1663–1673.

MacLeod, J. W., Lawrence, M. A., McConnell, M. M., Eskes, G. A., Klein, R. M., & Shore, D. I. (2010). Appraising the ant: Psychometric and theoretical considerations of the attention network test. *Neuropsychology*, *24*(5), 637.

Macmillan, N., & Creelman, C. (2005). *Detection theory, 2nd.* Lawrence Erlbaum Associates, New Jersey.

Macmillan, N. A., & Creelman, C. D. (1990). Response bias: Characteristics of detection theory, threshold theory, and" nonparametric" indexes. *Psychological bulletin*, *107*(3), 401.

Mahoney, S., Roth, E., Steinke, K., Pfautz, J., Wu, C., & Farry, M. (2010). A cognitive task analysis for cyber situational awareness. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 54, pp. 279–283).

Mancuso, V. F., Minotra, D., Giacobe, N., McNeese, M., & Tyworth, M. (2012). idsnets: An experimental platform to study situation awareness for intrusion detection analysts. In *2012 ieee international multi-disciplinary conference on cognitive methods in situation awareness and decision support* (pp. 73–79).

Mancuso, V. F., Strang, A. J., Funke, G. J., & Finomore, V. S. (2014). Human factors of cyber attacks a framework for human-centered research. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 58, pp. 437–441).

McConnell, M. M., & Shore, D. I. (2011). Mixing measures: testing an assumption of the attention network test. *Attention, Perception, & Psychophysics*, *73*(4), 1096–1107.

McIntire, L., McKinley, R. A., McIntire, J., Goodyear, C., & Nelson, J. (2013). Eye metrics: An alternative vigilance detector for military operators. *Military Psychology*, *25*(5), 502–513.

McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? on the influence of mind wandering and executive attention. *Journal of experimental psychology: general*, *141*(2), 302.

Medina, D., & Barraza, P. (2019). Efficiency of attentional networks in musicians and non-musicians. *Heliyon*, *5*(3), e01315.

Meyerhoff, H. S., & Papenmeier, F. (2020). Individual differences in visual attention: A short, reliable, open source, and multilingual test of multiple object tracking in psychopy. *PsyArXiv. June*, *10*.

Mickes, L., Wais, P. E., & Wixted, J. T. (2009). Recollection is a continuous process: Implications for dual-process theories of recognition memory. *Psychological science*, *20*(4), 509–515.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, *63*(2), 81.

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions four general conclusions. *Current directions in psychological science*, *21*(1), 8–14.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive psychology*, *41*(1), 49–100.

Monsell, S. (2003). Task switching. *Trends in cognitive sciences*, *7*(3), 134–140.

Motowildo, S. J., Borman, W. C., & Schmit, M. J. (1997). A theory of individual differences in task and contextual performance. *Human performance*, *10*(2), 71–83.

Newhouse, W., Keith, S., Scribner, B., & Witte, G. (2017). National initiative for cybersecurity education (nice) cybersecurity workforce framework. *NIST Special Publication*, *800*(2017), 181.

Northcutt, S. (2014, October). *Breeches happen: Be prepared* (Tech. Rep.). Sansinstitute.org: SANS Institute.

Northcutt, S. (2016). *Sec lab: Security products* (Tech. Rep.). www.sans.edu: SANS Institute.

Nunnally, J. C. (1978). An overview of psychological measurement. *Clinical diagnosis of mental disorders*, 97–146.

Paul, C. L., & Whitley, K. (2013). A taxonomy of cyber awareness questions for the user-centered design of cyber situation awareness. In *International conference on human aspects of information security, privacy, and trust* (pp. 145–154).

Petersen, R., Santos, D., Smith, M., & Witte, G. (2020). *Workforce framework for cybersecurity (nice framework)* (Tech. Rep.). National Institute of Standards and Technology.

Pettigrew, C., & Martin, R. C. (2016). The role of working memory capacity and interference resolution mechanisms in task switching. *The Quarterly Journal of Experimental Psychology*, *69*(12), 2431–2451.

Ponemon. (2020). *The 2020 devo soc performance report: A tale of two socs* (Tech. Rep.). ponemoninstitute.org: Ponemon Institute.

Posner, M. I. (1990). Hierarchical distributed networks in the neuropsychology of selective attention. *Cognitive neuropsychology and neurolinguistics: Advances in models of cognitive function and impairment*, 187–210.

Posner, M. I., & DiGirolamo, G. J. (1998). 1 8 executive attention: Conflict, target detection, and cognitive control.

Rajivan, P., Champion, M., Cooke, N. J., Jariwala, S., Dube, G., & Buchanan, V. (2013). Effects of teamwork versus group work on signal detection in cyber defense teams. In *International conference on augmented cognition* (pp. 172–180).

Rey-Mermet, A., Gade, M., Souza, A. S., von Bastian, C. C., & Oberauer, K. (2019). Is executive control related to working memory capacity and fluid intelligence? *Journal of Experimental Psychology: General*.

Riley, M., Elgin, B., Lawrence, D., & Matlack, C. (2014, March). *Missed alarms and 40 million stolen credit card numbers: How target blew it* (Tech. Rep.). www.bloomberg.com: Bloomberg.

Robertson, I. T., & Kandola, R. S. (1982). Work sample tests: Validity, adverse impact and applicant reaction. *Journal of occupational Psychology*, *55*(3), 171–183.

Roca, J., Crundall, D., Moreno-Ríos, S., Castro, C., & Lupiáñez, J. (2013). The influence of differences in the functioning of the neurocognitive attentional networks on drivers' performance. *Accident Analysis & Prevention*, *50*, 1193–1206.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of experimental psychology: General*, *124*(2), 207.

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.

Sauce, B., & Matzel, L. D. (2013). The causes of variation in learning and behavior: why individual differences matter. *Frontiers in psychology*, *4*, 395.

Sawaki, R., & Luck, S. J. (2010). Capture versus suppression of attention by salient singletons: Electrophysiological evidence for an automatic attend-to-me signal. *Attention, Perception, & Psychophysics*, *72*(6), 1455–1470.

Sawyer, B. D., Finomore, V. S., Funke, G. J., Matthews, G., Mancuso, V., Funke, M., … Hancock, P. A. (2015). Cyber vigilance. *American Intelligence Journal*, *32*(2), 151–159.

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: occupational attainment and job performance. *Journal of personality and social psychology*, *86*(1), 162.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological bulletin*, *124*(2), 262.

Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of applied psychology*, *71*(3), 432.

Shackleford, D. (2012, June). *When breaches happen: Top five questions to prepare for* (Tech. Rep.). Sansinstitute.org: SANS Institute.

Shanks, W. (2015). *Enhancing intrusion analysis through data visualization* (Tech. Rep.). Sansinstitute.org: SANS Institute.

Shipstead, Z., Harrison, T. L., & Engle, R. W. (2015). Working memory capacity and the scope and control of attention. *Attention, Perception, & Psychophysics*, *77*(6), 1863–1880.

Silva, A. R., McClain, J. T., Anderson, B. R., Nauer, K. S., Abbott, R., & Forsythe, J. C. (2014). *Factors impacting performance in competitive cyber exercises* (Tech. Rep.). Sandia National Lab.(SNL-NM), Albuquerque, NM (United States).

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information processing & management*, *45*(4), 427–437.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior research methods, instruments, & computers*, *31*(1), 137–149.

Stevens-Adams, S., Carbajal, A., Silva, A., Nauer, K., Anderson, B., Reed, T., & Forsythe, C. (2013). Enhanced training for cyber situational awareness. In *Foundations of augmented cognition* (pp. 90–99). Springer.

Svenmarck, P. (2020). –recruitment, selection and training of it/cyber personnel. *Human Systems Integration Approach to Cyber Security*, 25.

Theeuwes, J., & Burger, R. (1998). Attentional control during visual search: the effect of irrelevant singletons. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(5), 1342.

Trippe, D. M., Moriarty, K. O., Russell, T. L., Carretta, T. R., & Beatty, A. S. (2014). Development of a cyber/information technology knowledge test for military enlisted technical training qualification. *Military Psychology*, *26*(3), 182.

Tyworth, M., Giacobe, N. A., & Mancuso, V. (2012). Cyber situation awareness as distributed socio-cognitive work. In *Cyber sensing 2012* (Vol. 8408, p. 84080F).

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychological review*, *114*(1), 104.

Vandierendonck, A., Liefooghe, B., & Verbruggen, F. (2010). Task switching: interplay of reconfiguration and interference control. *Psychological bulletin*, *136*(4), 601.

Van Zandt, T. (2000). Roc curves and confidence judgments in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(3), 582.

Vieane, A., Funke, G., Greenlee, E., Mancuso, V., Borghetti, B., Miller, B., … Boehm-Davis, D. (2017). Task interruptions undermine cyber defense. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 61, pp. 375–379).

Vieane, A., Funke, G., Mancuso, V., Greenlee, E., Dye, G., Borghetti, B., … Brown, R. (2016). Coordinated displays to assist cyber defenders. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 60, pp. 344–348).

Waring, J. D., Chong, H., Wolk, D. A., & Budson, A. E. (2008). Preserved metamemorial ability in patients with mild alzheimer's disease: Shifting response bias. *Brain and Cognition*, *66*(1), 32–39.

Weaver, B., Bédard, M., McAuliffe, J., & Parkkari, M. (2009). Using the attention network test to predict driving test scores. *Accident Analysis & Prevention*, *41*(1), 76–83.

Wickens, C. D. (2008). Situation awareness: Review of mica endsley's 1995 articles on situation awareness theory and measurement. *Human factors*, *50*(3), 397–403.

Wickens, C. D., Gutzwiller, R. S., & Santamaria, A. (2015). Discrete task switching in overload: A meta-analyses and a model. *International Journal of Human-Computer Studies*, *79*, 79–84.

Wixted, J. T., Mickes, L., Dunn, J. C., Clark, S. E., & Wells, W. (2016). Estimating the reliability of eyewitness identifications from police lineups. *Proceedings of the National Academy of Sciences*, *113*(2), 304–309.

Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005). Rare items often missed in visual searches. *Nature*, *435*(7041), 439–440.

Woods, D. L., Kishiyama, M. M., Yund, E. W., Herron, T. J., Edwards, B., Poliva, O., … Reed, B. (2011). Improving digit span assessment of short-term verbal memory. *Journal of clinical and experimental neuropsychology*, *33*(1), 101–111.

Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive psychology*, *24*(3), 295–340.

Yantis, S., & Johnston, J. C. (1990). On the locus of visual selection: evidence from focused attention tasks. *Journal of experimental psychology: Human perception and performance*, *16*(1), 135.

Young, M. S., & Stanton, N. A. (2001). Mental workload: theory, measurement, and application. *International encyclopedia of ergonomics and human factors*, *1*, 507–509.

Yuan, H., Li, S., Rusconi, P., & Aljaffan, N. (2017). When eye-tracking meets cognitive modeling: applications to cyber security systems. In *International conference on human aspects of information security, privacy, and trust* (pp. 251–264).

Zhong, C. (2016). *A cognitive process tracing approach to cybersecurity data triage operations automation*. Unpublished doctoral dissertation, The Pennsylvania State University.

APPENDIX A. CONSENT FORMS

Purdue IRB Protocol number: 1703018965 - Expires: 08-MAY-2022

Key Information

RESEARCH PARTICIPANT CONSENT FORM

Neuro-Cyber Study

Dr. Marcus Rogers, Professor

Computer and Information Technology Kelly Cole

COT

Purdue University

Please take time to review this information carefully. This is a research study. Your participation in this study is voluntary which means that you may choose not to participate at any time without penalty or loss of benefits to which you are otherwise entitled. You may ask questions to the researchers about the study whenever you would like. If you decide to take part in the study, you will be asked to sign this form, be sure you understand what you will do and any possible risks or benefit. The purpose of this study is to identify cognitive abilities that would aid in cyber defense. We think this will take you 60 minutes, in one session, with two- minute breaks in-between. The harm or risk level for this study is no greater than everyday which is no greater than what you would encounter in daily life or during the performance of routine physical or psychological exams or tests. You will be compensated fifteen dollars.

What is the purpose of this study? The purpose of this study is to identify some of the mental abilities that may be of importance when defending a computer network so that so that educators, students, businesses and corporations will gain information that could improve how and who defends the network. Knowing what aspects of attention are important, in cyber defense tasks/environments, could help to better identify those who would work well in this field. You have been invited to participate in this study because COT/Purdue University has identified you as being enrolled in cyber courses. We are seeking to enroll up to 35 participants for this study.

What will I do if I choose to be in this study? If you decide to participate in this study, first you will be asked demographic questions such as, years of experience in Network/Cyber Security, age, gender and education level, followed by three neuro-psych tests. All the tests will be administered through a computer. You will enter an assigned ID (i.e., your name will not be recorded) and take each cognitive test starting with a judgment task where you are to identify the number presented to you as odd or even/greater-than or less- than, a digit span, an attention task, and lastly, finishing with the cyber task which includes classifying attacks in a simulated network environment.

How long will I be in the study? We think this will take you 60 minutes, in one session, with two-minute breaks in- between.

What are the possible risks or discomforts? The harm or risk level for this study is no greater than everyday which is no greater than what you would encounter in daily life or during the performance of routine physical or psychological exams or tests. A cognitive fatigue risk is present. We will offer breaks in-between all tests.

Are there any potential benefits? There are no direct benefits but your participation will benefit cyber defense research and society.

Will information about me and my participation be kept confidential? The data collected during the cognitive tests includes your reaction times and error rate. The data collected during the cyber simulation includes your completion time and error rate. Your name will not be recorded. This de-identified data will be stored in an excel sheet for data analysis-analyzed by the research team, and password-locked within in the PI's computer. All consent forms will be locked inside a locker in Knoy Hall-Cyber Forensics Lab. Breach of confidentiality is always a risk with data, but we will take precautions to minimize this risk as described in this confidentiality section. The project's research records may be reviewed by departments at Purdue University responsible for regulatory and research oversight.

What are my rights if I take part in this study? Your participation in this study is voluntary. You may choose not to participate or, if you agree to participate, you can withdraw your participation at any time without penalty or loss of benefits to which you are otherwise entitled. If you decide to withdraw from the study, the data that was collected will be deleted. Your decision to participate or not will have no effect on your relationship with Purdue University.

Who can I contact if I have questions about the study? If you have questions, comments or concerns about this research project, you can talk to one of the researchers. Please contact Kelly Cole at colek@purdue.edu or Dr. Marc Rogers at rogersmk@purdue.edu. If you have questions about your rights while taking part in the study or have concerns about the treatment of research participants, please call the Human Research Protection Program at (765) 494- 5942, email (irb@purdue.edu) or write to:

Human Research Protection Program - Purdue University
Ernest C. Young Hall, Room 1032
155 S. Grant St.,
West Lafayette, IN 47907-2114

To report anonymously via Purdue’s Hotline see www.purdue.edu/hotline

Documentation of Informed Consent

I have had the opportunity to read this consent form and have the research study explained. I have had the opportunity to ask questions about the research study, and my questions have been answered. I am prepared to participate in the research study described above. I will be offered a copy of this consent form after I sign it.

Participant’s Signature Here:

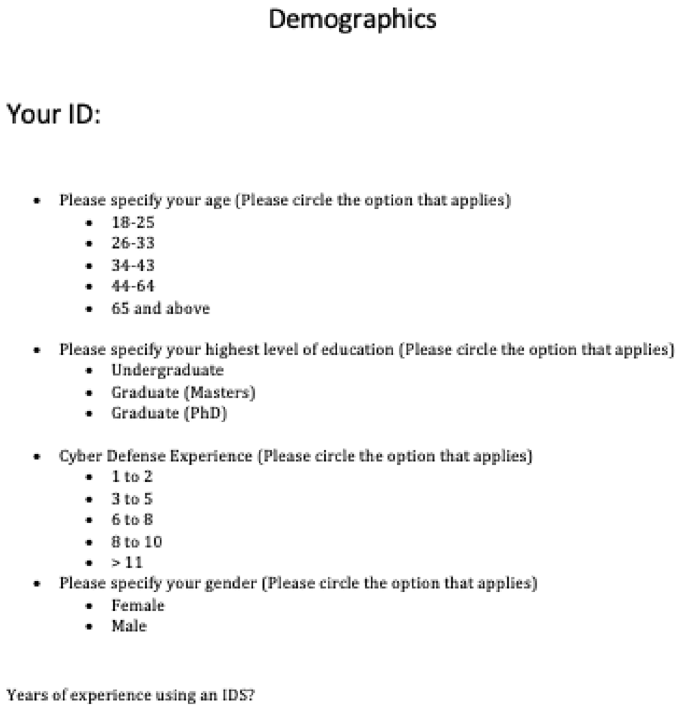
Participant’s Name Here:

Researcher’s Signature Here: Kelly Anne Cole

Date Here:

APPENDIX B. DEMOGRAPHIC QUESTIONNAIRE

The demographic questionnaire for the current study can be seen in Figure, B.1.



Demographics

Your ID:

- Please specify your age (Please circle the option that applies)
  - 18-25
  - 26-33
  - 34-43
  - 44-64
  - 65 and above

- Please specify your highest level of education (Please circle the option that applies)
  - Undergraduate
  - Graduate (Masters)
  - Graduate (PhD)

- Cyber Defense Experience (Please circle the option that applies)
  - 1 to 2
  - 3 to 5
  - 6 to 8
  - 8 to 10
  - > 11
- Please specify your gender (Please circle the option that applies)
  - Female
  - Male

Years of experience using an IDS?

Figure B.1. Demographic Questionnaire Administered to All Participants

APPENDIX C. CORRELATIONS

The correlations for all variables used in the study can be seen in Table, C.1.

Table C.1. *Correlation Table*

Variable		1	2	3	4	5	6	7	8	9	10	11	12	13
1. Hit Rate	n	–												
	Pearson's r	–												
	p-value	–												
	Spearman's rho	–												
	p-value	–												
2. FP Rate	n	16	–											
	Pearson's r	0.498	–											
	p-value	0.049	–											
	Spearman's rho	0.439	–											
	p-value	0.089	–											
3. CIAT Accuracy	n	16	16	–										
	Pearson's r	-0.437	-0.997	–										
	p-value	0.091	< .001	–										
	Spearman's rho	-0.394	-0.995	–										
	p-value	0.131	< .001	–										
4. Tool-Use	n	15	15	15	–									
	Pearson's r	-0.399	-0.546	0.540	–									
	p-value	0.141	0.035	0.038	–									
	Spearman's rho	-0.193	-0.489	0.494	–									
	p-value	0.490	0.064	0.061	–									
5. WM	n	16	16	16	15	–								
	Pearson's r	0.133	0.101	-0.101	0.115	–								
	p-value	0.624	0.710	0.710	0.682	–								
	Spearman's rho	0.134	0.146	-0.199	0.152	–								
	p-value	0.621	0.590	0.460	0.587	–								
6. WM	n	16	16	16	15	16	–							
	Pearson's r	0.509	0.229	-0.200	0.069	0.633	–							
	p-value	0.044	0.394	0.458	0.808	0.008	–							
	Spearman's rho	0.287	0.063	-0.095	0.029	0.686	–							
	p-value	0.281	0.817	0.727	0.918	0.003	–							
7. EF	n	16	16	16	15	16	16	–						
	Pearson's r	0.104	-0.017	0.033	0.060	-0.380	-0.252	–						
	p-value	0.702	0.950	0.902	0.833	0.147	0.346	–						
	Spearman's rho	0.062	0.018	-0.004	-0.113	-0.567	-0.352	–						
	p-value	0.819	0.948	0.987	0.689	0.022	0.181	–						
8. Orienting	n	16	16	16	15	16	16	16	–					
	Pearson's r	0.290	0.645	-0.654	-0.691	0.396	0.265	-0.024	–					
	p-value	0.277	0.007	0.006	0.004	0.128	0.322	0.929	–					
	Spearman's rho	0.204	0.617	-0.652	-0.742	0.337	0.306	0.024	–					
	p-value	0.449	0.011	0.006	0.002	0.202	0.248	0.935	–					
9. Alerting	n	16	16	16	15	16	16	16	16	–				
	Pearson's r	-0.734	-0.274	0.233	0.105	-0.388	-0.633	0.073	-0.190	–				
	p-value	0.001	0.304	0.384	0.710	0.137	0.009	0.787	0.482	–				
	Spearman's rho	-0.524	-0.108	0.112	0.154	-0.357	-0.424	0.191	-0.009	–				
	p-value	0.037	0.691	0.680	0.584	0.175	0.102	0.477	0.978	–				
10. SwitchCost(RT)	n	15	15	15	15	15	15	15	15	15	–			
	Pearson's r	0.085	0.602	-0.599	-0.259	-0.090	-0.119	0.209	0.463	0.714	–			
	p-value	0.763	0.018	0.018	0.350	0.750	0.673	0.455	0.082	0.003	–			
	Spearman's rho	0.077	0.526	-0.504	-0.247	-0.089	-0.147	0.093	0.446	0.682	–			
	p-value	0.785	0.044	0.055	0.375	0.751	0.600	0.743	0.097	0.007	–			
11. Spontswitrate	n	16	16	16	15	16	16	16	16	16	15	–		
	Pearson's r	0.336	-0.099	0.121	-0.138	0.508	0.467	-0.372	0.154	-0.588	-0.620	–		
	p-value	0.204	0.716	0.656	0.624	0.044	0.069	0.156	0.569	0.017	0.014	–		
	Spearman's rho	0.314	-0.121	0.087	0.006	0.508	0.499	-0.400	0.177	-0.580	-0.482	–		
	p-value	0.237	0.656	0.748	0.982	0.044	0.049	0.125	0.512	0.019	0.069	–		
12. SwitchingRT	n	15	15	15	15	15	15	15	15	15	15	15	–	
	Pearson's r	0.048	0.626	-0.620	-0.070	-0.002	-0.146	0.006	0.218	0.693	0.842	-0.520	–	
	p-value	0.865	0.013	0.014	0.805	0.994	0.604	0.983	0.436	0.004	< .001	0.047	–	
	Spearman's rho	0.039	0.624	-0.617	-0.023	0.026	-0.099	0.093	0.329	0.671	0.850	-0.470	–	
	p-value	0.891	0.013	0.014	0.934	0.928	0.726	0.743	0.232	0.008	< .001	0.077	–	
13. ambiSwitCosts	n	14	14	14	14	14	14	14	14	14	14	14	14	–
	Pearson's r	-0.751	-0.161	0.118	0.079	0.091	-0.026	0.075	0.130	0.510	0.330	-0.278	0.037	–
	p-value	0.002	0.584	0.688	0.788	0.757	0.930	0.800	0.659	0.062	0.249	0.337	0.899	–
	Spearman's rho	-0.669	-0.316	0.308	0.013	-0.198	-0.062	0.288	0.020	0.495	0.279	-0.294	0.051	–
	p-value	0.009	0.271	0.283	0.964	0.498	0.834	0.318	0.952	0.075	0.333	0.308	0.868	–