TRAC-M-TR-13-063
1 September, 2013

Understanding Optimal Decision-making in Wargaming

TRADOC Analysis Center - Monterey
700 Dyer Road
Monterey, California 93943-0692

This study cost the Department of Defense approximately $162,000 expended by TRAC in Fiscal Years 13.
Prepared on 20131029
TRAC Project Code # 060038

This page intentionally left blank.

TRAC-M-TR-13-063
1 September, 2013

Understanding Optimal Decision-making in Wargaming

MAJ Peter Nesbitt
Dr. Quinn Kennedy
LTC Jonathan Alt
Dr. Ji Hyun Yang
Dr. Ron Fricker
Dr. Jeff Appleget
Mr. Jesse Huston
MAJ Scott Patton
Mr. Lee Whitaker

TRADOC Analysis Center - Monterey
700 Dyer Road
Monterey, California 93943-0692

PREPARED BY:

MAJ Peter A. Nesbitt
MAJ, AR
TRAC-MTRY

APPROVED BY:

Christopher M. Smith
LTC, FA
Director, TRAC-MTRY

This page intentionally left blank.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01/09/2013	Technical Report	APR 2013 - SEP 2013

4. TITLE AND SUBTITLE
Understanding Optimal Decision-making in Wargaming

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
MAJ Peter Nesbitt
Dr. Quinn Kennedy
LTC Jonathan Alt
Dr. Ron Fricker
Mr. Lee Whitaker
Dr. Ji Hyun Yang
Dr. Jeff Appleget
Mr. Jesse Huston
MAJ Scott Patton

5d. PROJECT NUMBER
060038

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
TRADOC Research Analysis Center, Monterey, CA
Naval Postgraduate School, Monterey, CA

8. PERFORMING ORGANIZATION REPORT NUMBER
TRAC-M-TR-13-063

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Dr. Virginia Pasour, Army Research Office (ARO)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The research aims to gain insight into optimal wargaming decision making mechanisms using neurophysiological measures by investigating whether brain activation and visual scan patterns predict attention, perception, and/or decision-making errors through human-in-the-loop wargaming simulation experiments. We investigate whether brain activity and visual scan patterns can explain optimal wargaming decision making and its development with a within-person design; ie, the transition from exploring the environment to exploiting the environment. In this report, we first describe ongoing research that uses neurophysiological predictors in two military decision making tasks that tap reinforcement learning and cognitive flexibility.

15. SUBJECT TERMS
optimal decision-making, electroencephalography, neurophysiological predictors, wargames, eye tracking, military decision-making, cognitive factors, regret, reinforcement learning, sequential detection methods, Iowa Gambling Task, Wisconsin Card Sorting Task

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Peter Nesbitt
					19b. TELEPHONE NUMBER *(Include area code)*
U	U	U	SAR	110	(831) 656-7575

This page intentionally left blank.

TABLE OF CONTENTS

REPORT DOCUMENTATION iii

LIST OF FIGURES viii

LIST OF TABLES ix

EXECUTIVE SUMMARY x

LIST OF ACRONYMS AND ABBREVIATIONS xi

ACKNOWLEDGMENTS xii

1. INTRODUCTION 1
1.1. Support to ARO and TRAC Research Objectives 5

2. BACKGROUND 6
2.1. Military decision-making 6
2.2. Neurophysiological Measures 9
2.2.1. Electroencephalography (EEG) 9
2.2.2. Eye Tracking 11
2.3. Cognitive Factors in Underlying Optimal Decision 13
2.3.1. Iowa Gambling Test (IGT) 13
2.3.2. Wisconsin Card Sorting Test (WCST) 16
2.4. Mechanics of Optimal decision-making 18
2.4.1. Measure of Performance: Regret 18
2.4.2. Reinforcement Learning 20
2.5. Sequential Detection Methods for Detecting Exploration-Exploitation Mode Changes 22
2.5.1. Defining Some Sequential Methods 22

3. STUDY 1 METHODS 26
3.1. Experimental Design 26
3.1.1. Population of Interest 26
3.1.2. Decision-making Tests 26
3.1.3. Surveys 29
3.1.4. Covariate Measures 30
3.1.5. Equipment 30
3.1.6. Procedures 31
3.2. Analytic Methods 31

4. STUDY 1 PILOT DATA COLLECTION 33
4.1. Verification of Convoy and Map tasks 33
4.1.1. Verification of Convoy Task 33
4.1.2. Verification of Map Task 40

4.2. Verification of EEG and Eye Tracking System 41
4.2.1. Verification of EEG 41
4.2.2. Eyetracking 44

5. STUDY 2 THESIS **46**
5.1. Methodology 47
5.2. Neurophysiological Model of Tactical Decisions 48

6. CONCLUSION **49**
6.1. FY13 Progress 49
6.2. Initial Findings 50
6.3. Future Work 50

APPENDICES

APPENDIX A. Project Methodology **A-1**

APPENDIX B. Recruitment Advertisement **B-1**

APPENDIX C. IRB Approval of Study 1 **C-1**

APPENDIX D. Testing Procedure Tools **D-1**
D.1. Consent to Participate Form D-1
D.2. Procedure Notes D-3
D.3. Running the experiment: D-9
D.4. Demographic Survey Sheet D-11
D.5. Snellen Eye Acuity Test Chart D-12
D.6. Digit Span Memory Instructions D-13
D.7. Trails Making Test Instructions D-14
D.8. Trails Making Test A D-15
D.9. Trails Making Test B D-16
D.10.Convoy Task Penalty Script D-17
D.11.Post Task Survey Form D-23
D.12.Decision path chart for study 2 . . . . . D-25
D.13.Study Synopsis for RADML Doll visit D-26

REFERENCES **REF-1**

LIST OF FIGURES

1.1. Proposed hypothetical structure of decision-making. 5

2.1. Recognition primed decision-making. 6
2.2. Observe-Orient-Decide-Act (OODA) loop 7
2.3. Information processing model of human cognition. 8
2.4. Agent environment interaction. 8
2.5. Summary statistics for the original Iowa Gambling Task. 14
2.6. Image of score card from the original IGT. 15
2.7. Screen shot of the original Iowa Gambling Task. 15

3.1. Screen shot of the convoy task. 27
3.2. Description of graphics in map task. 28
3.3. Screen shot of the map task. 29
3.4. Data analysis methodology. 32

4.1. Route selection and latency by trial for pilot subject 1. 36
4.2. Route selection and latency by trial for pilot subject 3. 37
4.3. Regret for convoy task by pilot subject. 39
4.4. Regret comparison for convoy task. 39
4.5. Subject 4 frequency of cognitive state levels. 42
4.6. Subject 4 cognitive state levels by trial. 42
4.7. Subject 4 cognitive state levels by latency. 43
4.8. Gaze locations colored by route. 44
4.9. Eye tracking data from pilot sessions. 45

5.1. Examples of pop-up decision points. 46
5.2. Proposed hypothetical structure of decision-making. 48

A.1. Methodology flowchart for TRAC Project 638. . . . . A-1

B.1. Advertisement for recruitment of subjects. B-1

C.1. Naval Postgraduate School IRB approval of study 1. . . . . C-1

D.1. Consent to participate in research form. D-1
D.2. Demographic survey used before the Pilot Test. D-11
D.3. Snellen Eye Acuity Test Chart D-12
D.4. Digit Span Memory Test. D-13
D.5. Trails Making Test Instructions. D-14
D.6. Trails Making Test A. D-15
D.7. Trails Making Test B. D-16
D.8. Post Test Survey used to gain subject feedback on tasks. D-23
D.9. Synopsis provided to RADML Doll. D-26

LIST OF TABLES

2.1. WCST variables and their definitions. 17

3.1. Summary statistics for the convoy task. 27

4.1. Descriptive statistics of pilot subjects performance on convoy task. . . . 34
4.2. Descriptive statistics of pilot subjects 3 and 4 on convoy task. 38
4.3. Descriptive statistics of pilot subjects on map task. 40
4.4. Subject 4 EEG data during convoy task. 41

A.1. ARO Functions. A-2

D.1. Convoy task penalty script. D-17
D.2. Decision path chart for study 2. . . . . D-25

EXECUTIVE SUMMARY

Motivation
As the Army focuses on enhancing leader development and decision-making to improve the effectiveness of forces in combat, the importance of understanding how to effectively train decision-makers and how experienced decision-makers arrive at optimal or near optimal decisions has increased. Current understanding of how military decision-makers arrive at optimal decisions is not well understood and the measurement of decision-making performance lacks objectivity. The use of neurophysiological measures in human-in-the-loop wargames has the potential to fill this gap in knowledge and provide more objective measures of decision-making performance.

Purpose
This project's purpose is to investigate the role between neurophysiological indicators and optimal decision-making in the context of military scenarios as represented in human-in-the-loop wargaming simulation experiments. In this first year effort, we investigate optimal wargaming decision-making with a multi-pronged approach across two studies. Study 1 focuses on the development of optimal decision-making when all participants begin as naïve decision-makers. Specifically, study 1 attempts to identify the transition from exploring the environment as a naïve decision-maker to exploiting the environment as an experienced decision-maker via statistical and neurological measures. Study 2 examines wargaming decision-making in a dynamic and complex environment and will provide the opportunity to examine how different factors can contribute to optimal and non-optimal decision-making outcomes. In study 2 we will test our hypothetical structure of dynamic decision-making considering neural systems, gaze controls, and the world.

Army Relevancy and Military Application Areas
Objectively defining, measuring and developing a means to assess military optimal decision-making has potential to enhance training and refine procedures supporting more efficient learning and task accomplishment. Through the application of these statistical and neurophysiological models we endeavor to further neuromathematics and the understanding and modeling of decision making processes to more deeply understand the fundamentals of Soldier cognition. This project supports the Army's TRADOC Analysis Center's (TRAC) FY13 research requirements 1.2 - Agile Wargames, 2.6 - Mission Command Processes and decision-making, and 2.2 - Enhancing Subject Matter Expert (SME) Elicitation Techniques. The VA War Related Illness and Injury Study Center (WRIISC) is interested in this project to help identify PTSD and TBI. The results of this project also are of potential interest to the Neurophysiology Office and Simulations Office in ARL.

Summary of Current Status
We developed wargames that measure cognitive flexibility and reinforcement learning. We propose several statistical methods to objectively define and assess the transition to optimal decision-making. IRB approval for study 1 is granted. We successfully implemented and synchronized EEG to wargames and preliminary results of pilot data indicate validity of wargames and successful collection of neurophysiological markers.

LIST OF ACRONYMS AND ABBREVIATIONS

AI	artificial intelligence
ACT-R	adaptive control of thought–rational
ARO	Army Research Office
EEG	electroencephalography
EWMA	exponentially-weighted moving average
GOMS	goals, operators, methods, and selection rules
IGT	Iowa Gambling Task
MTRY	Monterey
NPS	Naval Postgraduate School
OODA	observe-orient-decide-act
RL	reinforcement learning
RPD	recognition primed decision
SARSA	state-action-reward-state-action
SOAR	state, operator and result
TRAC	TRADOC Analysis Center
TRADOC	Training and Doctrine Command
UCB	cpper control bound
VBS2	Virtual Battlespace 2
WCST	Wisconsin Card Sorting Task

ACKNOWLEDGMENTS

Dr. Paul Sajda - for advice on EEG systems, sharing of EEG techniques and of data.

Lt. Lee Sciarini, PhD, Assistant Professor at NPS - for demonstrating successful EEG calibration techniques and overall experience with the ABM EEG system.

Pilot Subjects - for volunteering their time, participation and feedback.

1. INTRODUCTION

As the Army focuses on enhancing leader development and decision-making to improve the effectiveness of forces in combat, the importance of understanding how to effectively train decision-makers and how experienced decision-makers arrive at optimal or near optimal decisions has increased. Army Chief of Staff, General Raymond T. Odierno, makes it clear that capable decision-making is important and necessary when he states "Future leaders must be adaptable, agile and able to operate in a threat environment that includes a combination of regular warfare, irregular warfare, terrorist activity and criminality." This is a descriptive account of what the future Army decision-makers should be, however, how a Soldier achieves adaptable and agile decision-making is poorly understood and defined. His priority on problem solving continues: "We have incredibly good leaders today, but we have to continue to develop them to address the many complex problems that I think we're going to face in the future." How does the Army meet this goal and develop a capability it doesn't understand? In following the Chief of Staff's lead, the US Army Maneuver Center of Excellence describes the desired effect of Soldiers making optimal decisions in Chapter 2 of the 2013 Maneuver Leader Development Strategy:

> CRITICAL THINKING AND PROBLEM SOLVING (CP)
> 2-8. Soldiers and leaders analyze and evaluate thinking, with a view to improving it. They solve complex problems by using experiences, training, education, critical questioning, convergent, critical, and creative thinking, and collaboration to develop solutions. Throughout their careers, Soldiers and leaders continue to analyze information and hone thinking skills while handling problems of increasing complexity. Select leaders develop strategic thinking skills necessary for assignments at the national level.

This qualitative description identifies the Army's need for leaders that can make optimal decisions. How can the Army meet its goals of finding, developing and providing for optimal decision-makers when the fundamental process of making an optimal decision is not well understood? One possibility here in the synthesis of behavioral and neurophysiological measurements of simulated wargaming decisions that are modeled via neuromathematics.

The Army Research Office (ARO), particularly the Biomathematics Program, identifies the importance of understanding the underlying mathematical fundamentals of optimal decision-making. The 2012 ARO Broad Agency Announcement states, "The ultimate goal of the Biomathematics Program focuses on adapting existing mathematics and creating new mathematical techniques to uncover fundamental relationships in biology, spanning different biological systems as well as multiple spatial and temporal scales. One area of special interest to the Program is neuromathematics, the mechanistic mathematical modeling of neural processes."

Wargames can inform how military decision-makers arrive at optimal decisions. Wargaming can be defined as a warfare model or simulation whose sequence of events affects, and is in turn affected by, the decisions made by players (Perla, 1990). The utility of modeling military

operations with wargames to prepare soldiers for future military operations is long recognized and demonstrated with examples from around the world in games such as chess, GO, Wei Hai and Chaturanga. The US Army utilizes wargaming as a technique to investigate decision-making under uncertainty. There are a number of beneficial outcomes from wargaming, one of which is a better understanding of the impact of decisions as a part of combat processes. However, using wargaming to understand decision-making processes and what factors affect these processes in real world, high cost decision-making is limited due to the complexities entailed in wargaming (i.e., multiple players, multiple decision points, multiple decision options at each decision point). Another limitation is that wargaming information is analyzed primarily with qualitative methods. If we know what factors affect decision-making and can measure these factors quantitatively, wargaming scenarios can be better geared towards optimizing training of military decision-making.

Observing decision-makers while playing wargames can give insight and inform on the underlying mathematical fundamentals of optimal decision-making. There is potential to understand how military decisions are made using wargames and their ability to simulate military scenarios. Neurophysiological measures, such as eye-tracking and electroencephalography (EEG), offer an objective and quantifiable method to understand how decision makers reason and arrive at decisions in high stress settings.

The past 25 years of research in eye movements have taught us that eye movements are not only evoked by the objects in visual scenes (i.e., external sensory signals) but also are responses to information about plans, goals, and probable sources of rewards or useful information. They are even responses to expectations about future events (Kowler, 2011). Thus, due to their close relations to visual attentional mechanisms, eye movements such as saccades can provide insight into cognitive processes that occur within the brain, such as language comprehension, mental imagery, and decision-making (Spivey et al., 2004).

Furthermore, recent technological advances have enabled non-intrusive implementation of eye-tracking feasible in complex and dynamic task environments. Research using eye-tracking technology has been very active and received significant attention in many domains. For example, eye scan behavior has been successfully utilized to detect expertise differences, such as in video gaming and driving (Mourant and Rockwell, 1972; Shapiro and Raymond, 1989). In the aviation domain, pilots exhibit different visual scanning patterns during various phases of flying under instrument flight rules (IFR) (Bellenkes et al., 1997; Katoh, 1997).

Of note, eye-tracking technology also has been applied to investigate decision-making in several situations, such as selecting a commercial product or flying a jet (Richardson and Spivey, 2004). However, one domain that has received significantly less attention with use of eye-tracking technology is military decision-making, i.e., understanding optimal decision-making in wargaming simulation. Eye movement data can be used to understand dominant factors enabling optimal decision-making in complex, dynamic, and uncertain warfare environments. For example, in an aviation decision-making study, optimal decision-making was positively correlated with attention allocation to problem relevant regions of interest (ROIs) after failure onset (Schriver et al., 2008). Specifically, expert pilots were faster to

notice, looked longer at, and responded faster to relevant cues when a failure was present than novice pilots. Eye-tracking parameters also have been found to predict errors in visual attention tasks, in which the frequency of long fixations (longer than 500 msec) was associated with the number of flight rule errors during a simulated flight (Van Orden et al., 2001). Another eye-tracking parameter, pupil diameter, is a reliable measure of cognitive workload (Van Orden et al., 2001); high cognitive load is associated with poorer decision-making (Gonzalez, 2005). These findings suggest that measuring changes in pupil diameter could be used to predict non optimal decision-making, and ultimately act as a warning signal before nonoptimal decisions can occur.

Additionally, physiological measures have the ability to determine fundamental components of decision-making processes. Event Related Potentials (ERP), the electrophysiological response to a sensory, cognitive, or motor stimulus (Luck, 2005), were observed in IED detection scenarios using VBS-2 training simulation (Skinner et al., 2010) and unique neural signatures of threat detection were identified. In some instances, physiological measures, including EEG, are more sensitive to initial changes in workload than performance-based measures (Brookings et al., 1996). For example, several different types of workload measurements were used to investigate workload on Air Force air traffic controllers in which traffic type and complexity was manipulated. EEG measurements were the only significant predictor of workload changes due to handling different types of traffic. For example, changes in heart rate predicted the type of decision made during a gambling task (Lee et al., 2010). Understanding physiological changes during wargaming could help determine why and when a non-optimal decision was made.

In our investigation into decision-making, we intend to assess methods of understanding how the process of assessing a state of the world and identifying the best action to take in a given state. Reinforcement learning, a sub-field of machine learning inspired by animal learning, provides one potential method of identifying the most appropriate action to take in a given situation (or state of the world) (Sutton and Barto, 1998; Watkins, 1989).

In this first year effort, we investigate optimal decision-making with a multi-pronged approach across two studies. Study 1 focuses on understanding and measuring how decision making expertise in a given decision situation is acquired by naïve decision makers. Specifically, study 1 attempts to identify the point at which decision makers transition from actions indicative of naive exploration of the decision environment to actions that indicate the decision maker is exploiting their knowledge to maximize their expected long term utility. Study 1 addresses the following issues for analysis:

1. How do we create a decision-making task that is military relevant, yet novel, so that all possible participants start as naïve decision-makers?
    (a) What cognitive characteristics or techniques should the task elicit to tap optimal wargaming decision-making?
    (b) How do we design the task(s) so that real-time neurophysiological measures of decision-making can be incorporated and synchronized with the participants' de-

cision behavior?

2. How do we determine when a wargamer demonstrates optimal decision-making?
   (a) How do we determine when a wargamer has transitioned from exploration to exploitation of the environment?
   (b) What are the neurophysiological markers that indicate the transition from exploration to exploitation?
   (c) How can we statistically model this transition?

Study 2 examines wargaming decision-making in a dynamic and complex environment and will provide the opportunity to examine how different factors can contribute to optimal and non-optimal decision-making outcomes. Using eye-tracking technology, different sources of decision-making errors can be classified based on gaze trajectory, and fixation frequency and duration. The addition of EEG measurements can help to determine the effects of internal disturbances (e.g., cognitive workload, confidence) and external disturbances (e.g., other stimuli) on decision-making processes.

Figure 1.1 shows our proposed hypothetical structure of dynamic decision-making considering neural systems, gaze controls, and the world. The schema control concept is adopted from (Land and Hayhoe, 2001), in which schema control refers to the command and control center of the brain. Errors related to decision-making processes can be modeled in the following four hierarchical levels. First, eye movement information can provide whether human operators looked at relevant information during wargaming scenarios. If foveal vision misses significant information (Level 1 errors or attention errors), it is obvious that optimal decision-making cannot be reached. Second, even when some important information is looked at, but not long enough (Level 2 errors) or due to internal/external disturbances (Level 3 errors), human operators cannot perceive the information correctly. Level 2 and Level 3 errors can be categorized as perception errors, where Bayesian modeling approach and EEG will be used to describe different types of perception errors. Finally, decision errors can appear even when no attention or perception errors are associated (Level 4 errors or decision errors). For example, decision outcomes can be non-optimal due to inherent bias (e.g., the decision is preset by schema control even before information has been scanned), within-subject differences or between subject differences.

Study 2 will address the questions:

1. How well does our hypothesized model of dynamic decision-making predict optimal and nonoptimal decisions?
2. Do military personnel who must rely on new technology (unmanned wingman) to make tactical decisions show different visual scan and EEG patterns prior to making their decisions than those relying on traditional means (live wingman)?

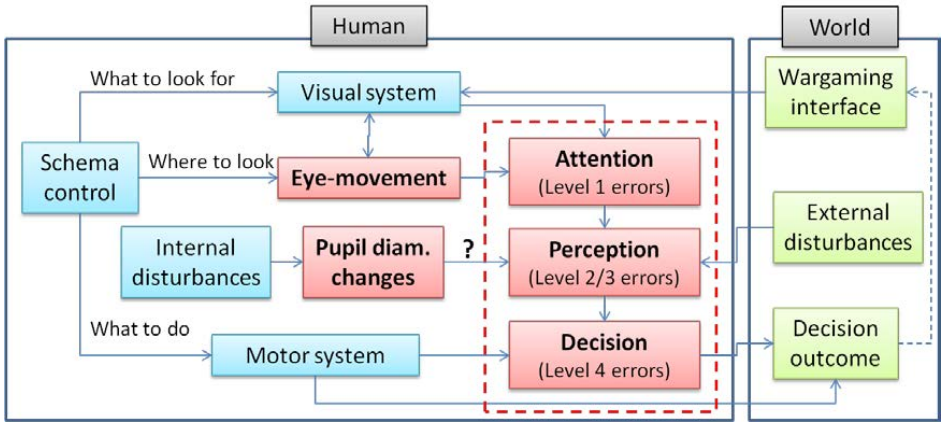


Figure 1.1: Proposed hypothetical structure of decision-making considering neural system, gaze control, and world.

In sum, through the synthesis of neurophysiological and behavioral decision measurements, the proposed research will extend current understanding of the development of optimal wargaming decision-making as well as test a hypothesized model of dynamic decision-making. Additionally, the implementation of novel statistical methods based on reinforcement learning theory, as described in Section 2.5 and 2.4, will provide greater insights into understanding how military personnel transition from naïve to optimal decision-making.

1.1. Support to ARO and TRAC Research Objectives

This project supports TRAC research requirements and ARO functions. This technical report covers the preparation, execution and results of the first year of this three year effort. It documents the investigation of the role between neurophysiological indicators and optimal decision-making in the context of military decision-making scenarios as represented in human-in-the-loop wargaming simulation experiments. This project supports TRAC research requirements as defined by the FY13 Research Plan. The specific requirements directly supported are 1.2 Agile Wargames, 2.6 Enhancing Subject Matter Expert (SME) Elicitation Techniques, and 2.2 Mission Command Processes and decision-making (Alt et al., 2013). ARO research interests are defined by the Army Research Office Functions (ARO, 2012). The project satisfies these research objectives by following the approved methodology, graphically represented in Figure A.1.

2. BACKGROUND

2.1. Military decision-making

Military decision-making has been studied extensively by the situational awareness community (Klein, 1993, 2008). In this section, we briefly outline some of the commonly used decision-making models. Klein developed the recognition primed decision (RPD) making model, a form of naturalistic decision-making, that posits that those with appropriate experience in a given situation learn to recognize the situation and identify the optimal action to take in that situation. This model implicitly relies on the concept of reinforcement learning as described in the literature on animal learning (Thorndike, 1911). Klein reported on multiple case studies on expert leadership from emergency responders and from the military (Klein, 1993, 2008). See Figure 2.1 for his diagram.

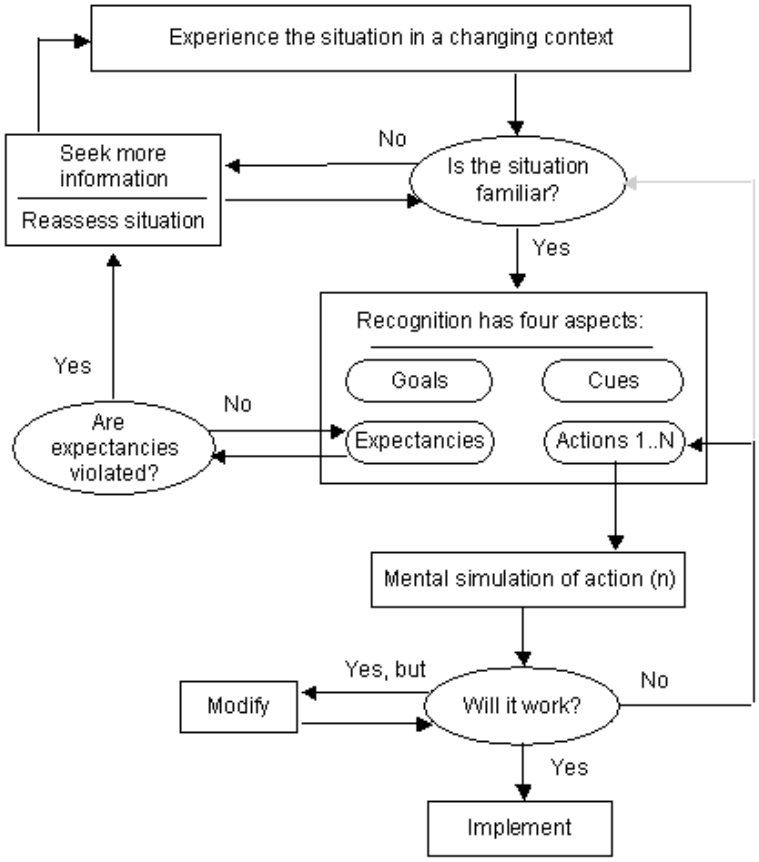


Figure 2.1: Recognition primed decision-making (Klein, 1993).

Boyd describes the Observe-Orient-Decide-Act (OODA) loop in the context of pilot engagements (Boyd et al., 2007). The OODA loop has been used to describe the process of military decision-making in a number of other contexts, and can be seen as a diagram in Figure 2.2.

The parallel between the OODA loop and Klein's RPD model are relatively straightforward to grasp.

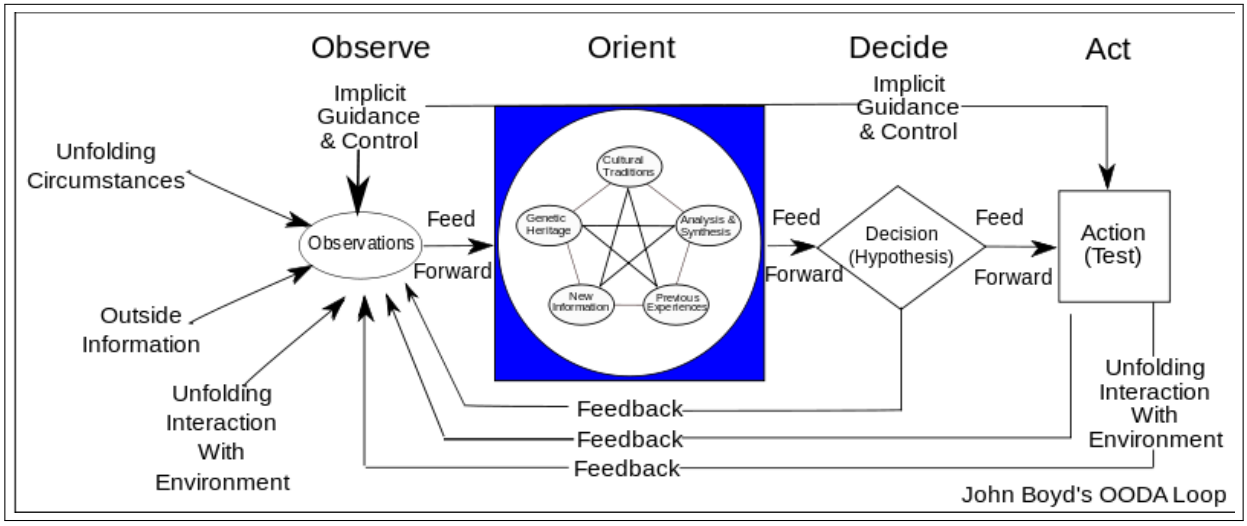


Figure 2.2: Observe-Orient-Decide-Act (OODA) loop in the context of pilot engagements (Boyd et al., 2007).

Wickens proposes an information processing model of cognition that fits nicely in the context of Klein's RPD model and with Boyd's OODA loop (Wickens and Carswell, 1997; Wickens and Hollands, 2000). In this model, Wickens abstracts cognition to perception, guided by attention, working and long-term memory, and some internal decision-making model that relies on perceived information and memory to identify the appropriate decision. This model is generally in line with Anderson and other models of cognition.

Finally, two models that focus on the decision-maker's internal state, rather than on the actual decision are the dynamic model of situation cognition (Miller and Shattuck, 2004) and situational awareness (Endsley, 1995). Shattuck and Miller propose the dynamic model of situated cognition as a means to explain how decision-making occurs (Miller and Shattuck, 2004). They use a paradigm of six lenses that describe how the decision-maker perceives the environment and identifies a decision state. Their model does not go into detail on how the decision-maker selects from the set of actions available in a given state. Endsley developed the concept of situational awareness and describes four levels of situational awareness (Endsley, 1995). Similar to Shattuck and Miller, Endsley does not develop the action selection mechanism, instead focusing on how the decision-maker develops a representation of the perceived state of the world.

The artificial intelligence community relies on human decision-making as an inspiration for devoting autonomous software agents. The general model of the agent environment interaction is similar to those models described previously (Russell and Norvig, 2010). The perception of information from the environment is accomplished through a sensor, the agent reasons about the appropriate action to take using some internal model, and the action is

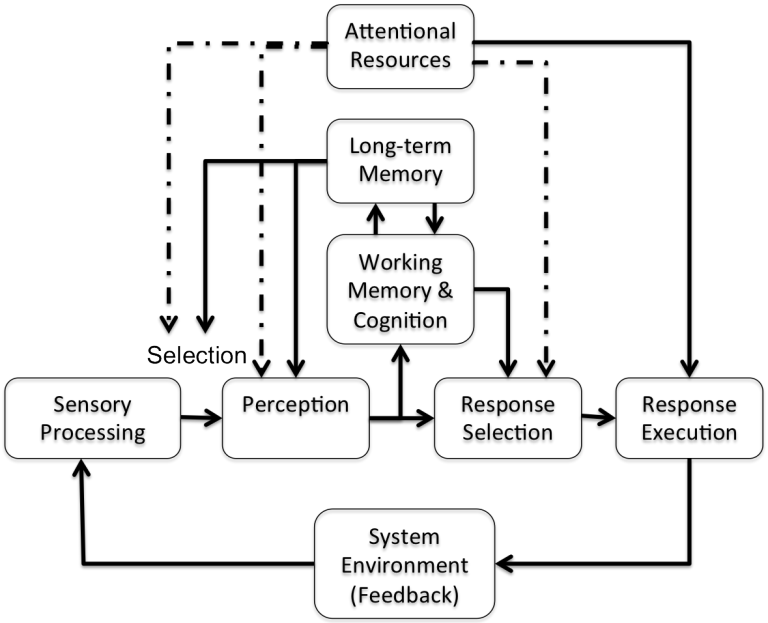


Figure 2.3: Information processing model of human cognition (Wickens and Hollands, 2000). Dotted lines from attentional resources intend to illustrate the many places selective attention interacts with the processing of information.

expressed in the environment in which the agent is operating. Those agent architectures that attempt to constrain the agent in the same manner that a human would are referred to as cognitive architectures, with adaptive control of thought–rational (ACT-R) and state, operator and result (SOAR) serving as prominent examples (Zacharias et al., 2008; Laird and Wray III, 2010). The internal reasoning employed within these systems can take the form of a rule based expert system, sometimes developed using a goals, operators, methods, and selection rules (GOMS) type of cognitive task analysis, or make use of learning systems that are given only some notion of a goal, a reward, a perception of state, and a set of available actions.

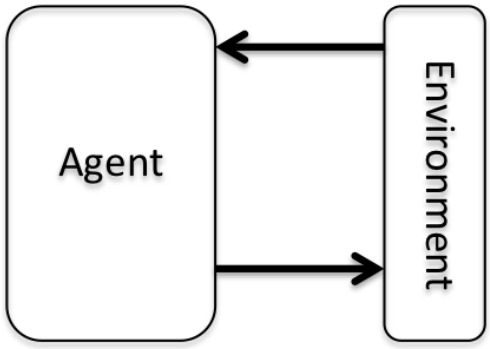


Figure 2.4: Agent environment interaction (Russell and Norvig, 2010).

The cognitive science community has begun to make use of of eye-tracking and EEG technology to study decision-making in a variety of contexts, however, the military decision-making research has not made extensive use of these technologies to study decision-making (Zacharias et al., 2008). The potential to gain additional understanding of factors affecting the quality of military decision-making exists through these technologies. The potential also exists to leverage these technologies to gain additional insights in applied areas that leverage techniques such as wargaming to inform analysis questions (force structure questions, course of action analysis, design of control systems). There are gaps in our understanding of military decision-making (Perla, 1990).

2.2. Neurophysiological Measures

The study of neurophysiological factors has the potential to fill these gaps. EEG has been used successfully to describe neurophysiological activity during decision-making (Gluth et al., 2013b,a), as well as to untap neurophysiological differences between experts and novices in a variety of tasks (Sherwin and Gaston, 2013; Herzmann and Curran, 2011; Ott, 2013), and even the development of expertise (Krigolson et al., 2009). Additionally, numerous studies indicate that visual scan data via eye tracking technology can provide valuable insights into participants cognitive strategies during real-world tasks; strategies that cannot be detected by behavioral performance alone (Kasarskis et al., 2001; Marshall, 2007; Sullivan et al., 2011; Yang et al., 2011; Van Orden et al., 2001; Cowden, 2012; Sullivan et al., 2011). Therefore, the combination of EEG and eye tracking technology provides a much finer-grained signature of the stages of development towards optimal decision-making than behavioral performance data. Not only do they provide neurophysiological measurements not tapped by behavioral data, but they also provide important information regarding participants cognitive state throughout the entire decision-making process in each trial, rather than just information regarding the final decision at the end of the trial.

Below, we provide a preliminary literature review targeted on how EEG and eye tracking technology can be used to understand underlying cognitive states in both dynamic and static decision-making tasks. Because experts typically demonstrate optimal decision-making in the domain of expertise compared to their less experienced counterparts (Ericsson, 2006), we also examine how these neurophysiological measures can tap expertise differences in decision-making as well characterize the development of expert decision-making.

2.2.1. Electroencephalography (EEG)

Electroencephalography (EEG) is the recording of voltage fluctuations resulting from ionic current flows within the neurons of the brain. We use EEG data as neurological indicators of decision-making.

2.2.1.1. EEG sensitive to different cognitive states

EEG has been used reliably to monitor cognitive state fluctuations during both dynamic, real world decision tasks and static, trial by trial decision-making tasks. Berka et al. (2007) demonstrated that EEG can provide an unobtrusive method for monitoring dynamic fluctuations in cognitive state (Berka et al., 2007). While Navy operators completed a three hour simulated missile execution scenario that entailed 8 phases and 26 tasks, their brain activity was measured via EEG in one second increments. EEG data was classified by level of engagement,workload, and drowsiness. The EEG engagement index captures information gathering, visual scanning, and sustained attention. EEG workload index measures executive functions such as working memory load, problem solving, information integration, and mental math. Results indicated, that as expected, engagement and workload measures levels were positively correlated with phases in which high activity was required by participants and negatively correlated with phases that required low activity by participants. Importantly, the engagement and workload measures were independent of each other. As we are using the same EEG system, these results indicate that we will be able to capture real time measures of cognitive workload and engagement.

Whereas Berka et al. (2007) were interested in creating global indices of cognitive state, other research has looked at the raw EEG signals to determine how brain activity correlates with real world decision-making, such as choosing a medication (Davis et al., 2011). EEG responses were divided into bands, with a focus on the alpha band (decreases in alpha band are associated with increased brain activity and cognitive workload, whereas increases are associated with a relaxed and focused cognitive state) (Cahn and Polich, 2006; Oakes et al., 2004). The decision scenarios contained three attributes (eg, cost, convenience, and quality), each of which varied in value (low, medium, high). Each scenario was presented twice. For each scenario, an optimal decision was calculated as the highest sum of the values from each individual attribute. Decision performance measures included errors (when participants chose an option that had a lower total value than the optimal decision, inconsistency (when a participant chose one option on the first presentation of that scenario and then chose a different option on the 2nd presentation of that scenario), and response time. As hypothesized, results indicated that the alpha band was positively correlated with errors and inconsistency, and to some extent, response time and negatively correlated with expressed preference. Thus, results demonstrate how individual differences in decision-making performance uniquely relate to changes in raw EEG activity.

Besides being able to measure dynamic changes in cognitive state, EEG measurements also can detect differences in cognitive state in trial by trial tasks that involve static images. Campbell et al. (2009) investigated intentional responses versus guesses and slips (such as accidentally hitting the wrong key). Participants completed a recognition task in which they completed a familiarization phase regarding the relationship between tank silhouettes, their names, and the key on the keyboard associated with that tank silhouette. Each tank could be classified by seven different types of visual features. Participants then were shown only the tank silhouette and were instructed to press the key associated with that tank. Results revealed that when all of the decision data was used, it appeared that participants rarely

used the visual features appropriately, suggesting that additional training would be needed. In contrast, when EEG-detected guesses and slips were removed from analyses, the opposite findings emerged: participants were using most visual features appropriately. The researchers demonstrate that that when guesses and slips are removed from the analyses, a very different picture of participant performance and understanding of the task emerges. In a follow up study, Campbell et al (2011) demonstrate that distinguishing between intentional responses from guesses and slips can predict the likelihood of future performance errors (Campbell et al., 2011). In particular, they found that whereas EEG-detected guesses and slips were less likely in future performance sessions, frequency of intentional errors remained consistent. Taken together, these results clearly demonstrate the utility of using EEG measurements to inform the training needs of a given individual.

2.2.1.2. EEG sensitive to expertise differences in decision-making

Evidence also is accumulating that EEG can untap neurophysiological differences between experts and novices in a wide range of tasks (Holroyd et al., 2009; Herzmann and Curran, 2011; Ott, 2013; Sherwin and Gaston, 2013). Most relevant to the current study is work by Krigolson et al. (2009), in which they investigated the role of the medial-frontal reinforcement learning system via EEG measurements in the development of perceptual expertise (Holroyd et al., 2009; Holroyd and Krigolson, 2007). Previous research has found that the medial frontal cortex plays a key role in reinforcement learning (Holroyd et al., 2005). Participants were asked to discriminate between two sets of learnable blob shapes, and a set of morph blobs with characteristics from both sets, receiving feedback after each response. As in our proposed task, participants were näive to the perceptual discrimination task, and therefore had to learn the discrimination rules through trial and error. Two components of EEG measurements were examined, the N250 and the error related negativity (ERN). The N250 has been shown to increase in amplitude when experts view objects in their domain of expertise (Scott et al., 2008). ERN can distinguish between correct and incorrect responses in speeded response time tasks (Gehring et al., 1993), as well between correct and incorrect feedback in trial-and-error learning tasks (Holroyd and Krigolson, 2007). Changes in N250 amplitude distinguished between high learners (those with response accuracy greater than 70%) and low learners (those with response accuracy less than 70%). For high learners, N250 amplitude increased as performance improved. However, for low learners, no such change in N250 amplitude occurred. These findings suggest that high learners, who learned to correctly identify blobs, were able to internally evaluate the consequences of their behavioral responses and thus benefit from reinforcement learning. More broadly speaking, the results indicated that the development of perceptual expertise relies on interactions between the posterior perceptual system and the reinforcement learning system involving medial-frontal cortex.

2.2.2. Eye Tracking

Eye-tracking technology provides nonintrusive devices to collect ocular data which makes it ideal to measuring visual scan patterns in both laboratory settings and real operational environments. Common visual scan measures collected from eye tracking are fixations (gazing at something for more than 70 milliseconds), saccades (rapid movements of eyes), dwell

duration (also called fixation duration; the interval between two successive saccades), and blink rate. Based off these metrics, different cognitive states and cognitive strategies can be detected. For example, opposing cognitive states, i.e., engaged vs. relaxed, normal vs. distracted, and fatigued vs. alert, were modeled from eye movement and pupil size (Marshall, 2007); other research demonstrates that eye-tracking measurements can be used to determine level of visual processing load (Van Orden et al., 2001). Along with others, our work with helicopter pilots has shown that visual scan pattern can be used to detect underlying cognitive strategies that the pilots may not even be aware of using (Di Nocera et al., 2007; Yang et al., 2011). Below we describe relevant findings regarding eye tracking and domain specific expertise, and how the combination of eye tracking technology combined with reinforcement learning theory can enable us to understand how visual scan patterns change as expertise is developed.

2.2.2.1. Eyetracking and expertise

Similar to EEG, eye tracking has been used reliably to detect expertise differences in a variety of tasks, such as driving, flying an aircraft, and even chess (Borowsky et al., 2010; Charness et al., 2001; Kasarskis et al., 2001; Sullivan et al., 2011). The recurrent finding in this literature is that experts tend to have a more efficient and effective visual scan pattern than the less expert. For example, expertise differences in visual scan pattern were clearly exhibited in a study in which expert and novice pilots performed a simulated landing (Kasarskis et al., 2001). Expert pilots had a targeted visual scan pattern that alternated between looking at the runway and the airspeed indicator (the most salient instrument during landing). In contrast, novice pilots showed a weak visual scan pattern between the runway and airspeed indicator and instead tended to make several consecutive fixations in the general area of runway and many horizontal saccades within the runway. Additionally, expert pilots showed a more efficient visual scan pattern by having significantly more fixations in general, lower average dwell time, more time spent looking at the runway aimpoint time, and greater number of airspeed indicator fixations. Importantly, better landings were associated with more fixations and shorter dwells, indicating that visual scan is correlated with actual landing performance.

In another dynamic task in which experienced and novice drivers completed a simulated driving scenario, experienced drivers were more likely to detect and respond to potential, unexpected hazards compared to novices. They also were more likely to look to the right of a T intersection, whereas novice drivers continued to look straight ahead (Borowsky et al., 2010). The above studies focused on dynamic tasks. Eyetracking also can detect expertise in a more static, trial by trial, environment. In a study examining visual scan patterns of expert and intermediate chess players, experts visual scan pattern indicated that experts are more efficient in encoding the information presented on the chess board (Charness et al., 2001). Specifically, expert chess players made larger saccades, fixated more on empty squares and salient pieces, and had about half as many fixations per trial than intermediate chess players. The finding that experts focus more on empty squares than intermediates suggests that experts use their domain-specific knowledge to encode chess configurations rather than individual pieces.

These studies demonstrate that under conditions that range from relatively static to fast paced scenarios, already established experts focus on the salient information, whether that be pertinent aircraft instruments, the most dangerous area of a T intersection, or the key chess pieces. The question remains as to how visual scan pattern changes as expertise develops. The recent advances in both eye tracking technology and theoretical development in reinforcement learning now enable this question to be addressed (Hayhoe and Ballard, 2005). Visual scan patterns are learned through reinforcement learning: the novice must learn what stimuli are important in the context of the task, where that stimuli typically are located, and eventually, to anticipate relevant stimuli by moving their gaze to the location of that stimuli prior to the event of interest. Saccades, movements in which the eye gaze moves from one fixation point to another, is highly sensitive to positive and negative feedback (Hayhoe and Ballard, 2005). Thus, reinforcement learning appears to play a key role in the development of expert visual scan.

2.3. Cognitive Factors in Underlying Optimal Decision

During the task development phase of the project, the team spent a lot of time discussing what exactly a soldier needs to make optimal decisions, as well as to transition from näive to optimal decision-making. Based on these discussion and substantiated by the literature (Yang et al., 2011), we determined that there are at least two key components: reinforcement learning, the ability to learn from trial and error, and cognitive flexibility, the ability to recognize when the rules have changed or that the current strategy no longer works. We determined that two common psychological tests that measure reinforcement learning and cognitive flexibility, the Iowa Gambling Task (IGT) (Bechara et al., 1994) and the Wisconsin Card Sorting Task (WCST) (Grant and Berg, 1948), could be modified into wargaming tasks. Additionally, the IGT and WCST are static, trial by trial decision tasks, wargaming adaptions of these tests allow for fine-grained investigation into the development of optimal decision-making. Below, we provide a brief overview of the IGT and WCST and research investigating the neurophysiological correlates of IGT and WCST decision performance.

2.3.1. Iowa Gambling Test (IGT)

The Iowa Gambling Test (IGT) is a psychological task developed at the University of Iowa in 1994 used to measure decision-making performance in the presence of uncertainty, known as a multi-armed bandit problem in the artificial intelligence and operations research literature. The original IGT was developed to identify behavior particular to patients with damage to the ventromedial prefrontal cortex (Bechara et al., 2005). The test was developed as a "neurophysiological task which simulates, in real time, personal real-life decision-making relative to the way it factors uncertainty of premises and outcomes, as well as reward and punishment" (Bechara et al., 1994). Damage to this brain region is associated with difficulty in learning from trial and error, particularly for personal decisions, and in choosing options with uncertain outcomes.

This test is useful to the ODM project for it causes the subjects to rely on their ability to develop an estimate of long term payoff for decision-making. There are many examples of IGT use in the professional literature, providing a foundation of published work to leverage for ODM. However, these studies apply the original IGT differently, leaving no single standard. This leaves an opportunity to develop an IGT for the ODM project that meets the needs of this project in a manner that is also consistent with the most appropriate previous applications of the test.

2.3.1.1. Materials and methods of the original IGT

For the original IGT, the subjects receive a loan of $2000 of play money and are asked to make a series of decisions to maximize the profit on the loan. Each decision entails selecting one card at a time from any of four available decks of cards. The subjects are told the game requires a long series of card selections, and then continue selecting cards until they are told to stop (the task is stopped after 100 card selections). All cards give money and some cards also require a penalty. The summary statistics for the original IGT are in Figure 2.5. It is expected that healthy subjects first go through a period of exploration in which they seek to determine which decks have the best long term payoffs (decks C and D). At some point the subject gains enough confidence in their estimate of the payoff for each alternative that they begin to make use of this perceived information to maximize their long-term reward by choosing the alternatives with the greatest estimate - this phase is referred to as the exploitation phase. Typical decision performance measures are net score, frequency with which each deck was selected, an advantageous selection bias (proportion of good decks selected minus the proportion of bad decks selected), and intermediate deck selection frequencies (deck selection frequencies within each block of 20 trials) (Bechara et al., 1994; Steingroever et al., 2013).

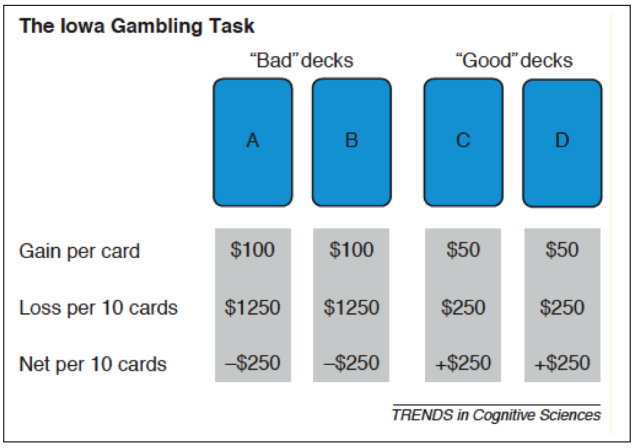


Figure 2.5: Summary statistics for the original Iowa Gambling Task (Bechara et al., 2005).

These summary statistics offer an understanding of the long term value of each deck. However, the subject only has access to the value of each deck by trial. An example of the by trial value of each deck is found in Figure 2.6.

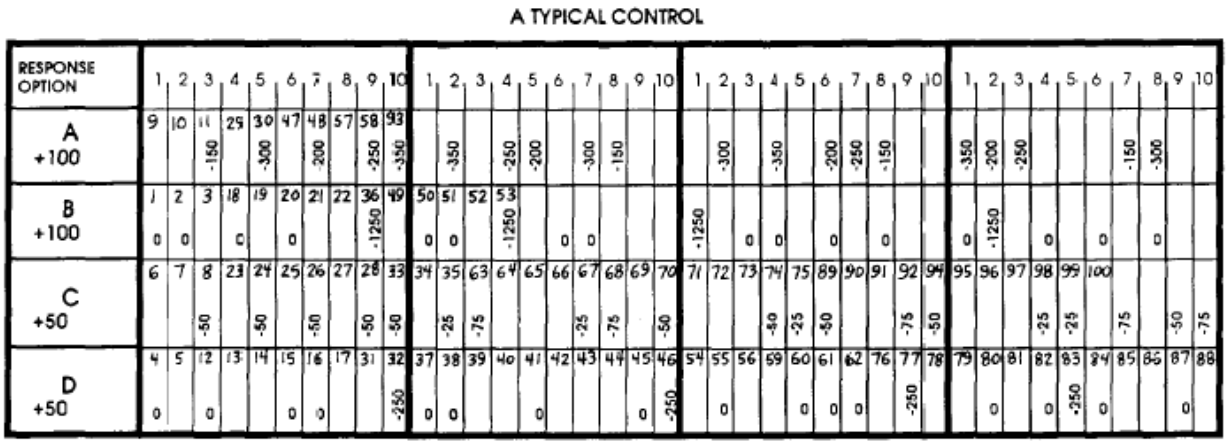


Figure 2.6: Image of score card from the original IGT, depicting the performance of a typical control subject. The hand written numbers are the order of which the subject drew each card. Bechara et al. (1994).

Nearly all applications of the IGT leave minimal visual difference between images representing the available options. The intent of similar looking options is to minimize the visual bias. This is demonstrated in a screen shot of the first IGT in Figure 2.7. The test purposely precludes any visual indicator of when the test will end.

Figure 2.7: Screen shot of the original Iowa Gambling Task by Bechara et al. (2005). The *A'* deck is black for it is the last deck chosen, winning $120 for the player. The smiley face on the left is further positive stimuli for choosing a deck that reurned a positive value. There are two meters on the top, one green and one red. The green is labeled *Cash Pile* and the red is labeled *Borrowed*. These meters reflect the cumulative player performance up to the current trial.

Although most studies of IGT have concentrated on its ability to distinguish between healthy and clinical populations, recent studies have focused on the variability in decision performance found among healthy participants (Steingroever et al., 2013; Worthy et al., 2012). Steingroever et al conducted a meta-analysis to determine if healthy participants consistently (1) select good decks to a greater extent than bad decks, and (2) go through a clear exploration to exploitation pattern. These decision patterns were not evident. Instead, healthy participants actually weigh frequency of losses (decks B and D) more so than longterm payoffs (decks C and D). Additionally, high levels of individual differences in decision patterns were revealed; for example, in some cases, individuals remained in the exploration phase; whereas others showed idiosyncratic exploitation patterns: continuously switching back and forth between Decks B and D (eg, DBDBDBDBDB pattern), sticking with one deck until receiving a loss (eg, BBBBBDDDBBBBDDDDD), or consistently sticking with one deck regardless of losses (eg, BBBBBBBBBBBBBBBBBBB). Thus, the high level of individual variability in decision patterns is important for two reasons: (1) it parallels real-world military decision-making in which there may be several different, equally viable, paths that lead to optimal decision-making, and (2) it demonstrates a clear need for more sophisticated modeling techniques that can account for inter-individual variability in decision patterns. Indeed, recent developments in reinforcement learning have been used to model IGT decision patterns (Worthy et al., 2012).

2.3.1.2. N-arm Bandit Problem

In the study of probability theory and reinforcement learning, the IGT is an example of a N-armed bandit problem. A n-armed bandit problem is the situation where a player is repeatedly faced with a choice among different options, or actions. After each choice the player receives a numerical reward chosen from a stationary probability distribution that depends on the action selected (Sutton and Barto, 1998). Most examples of the IGT do not rely on a probability distribution function. Instead each deck is a scripted, ordered set of cards with specified values as seen in Figure 2.6. The N refers to the number of options available to the player. The IGT generally offers only four options to the player, making $N = 4$. As the player makes choices, they develop an estimate of the expected value for future action. A *greedy* action is defined as the player selecting a choice for its highest estimate. If a player selects a greedy action, they are *exploiting* their current knowledge of the estimate of the choices. If the player selects a choice with a smaller estimate, they are *exploring* the other choices in order to improve their estimate. The player is simultaneously attempting to improve their estimate of the options and optimize this knowledge to maximize their score. Understanding the theory and leveraging published work on N-arm Bandits gives insight to available methods to understanding military decision-making (Audibert et al., 2007; Auer et al., 2002; Auer and Ortner, 2010).

2.3.2. Wisconsin Card Sorting Test (WCST)

The WCST taps the working memory, shifting and inhibition components of executive function and therefore is heavily reliant on prefrontal cortex, basal ganglia, and thalamus functioning (Huizinga and van der Molen, 2007; Konishi et al., 1999; Monchi et al., 2001). Par-

ticipants view 5 cards, one card displayed at the top center of the screen, the remaining four displayed across the bottom of the screen. Each card contains symbols that vary in number, shape, and color. Over several trials, participants try to figure out the matching rule that will correctly match the card on the top of the screen with one of the four cards at the bottom of the screen. Unbeknownst to the participant, the matching rule changes once the participant has 10 consecutive correct matches. For example, after 10 consecutive correct matches based on the color of the symbols, the matching rule could then change to the number or shape of the symbols. Thus, participants must not only learn and maintain in working memory the correct matching rule while inhibiting irrelevant stimuli, but also exhibit cognitive flexibility in detecting when the rule has changed (Grant and Berg, 1948; Huizinga and van der Molen, 2007). The task is completed when either the participant has successfully completed two rounds of each matching rule or until they have completed 128 trials. Table 2.1 demonstrates typical decision performance measures, which link the particular executive function component to the type of errors made, e.g. perseverative (failure to show cognitive flexibility) verses failure to maintain set (inability to retain current matching rule while inhibiting irrelevant stimuli).

Variable	**Definition**
Time	Time taken on each trial.
# trials	Total number of trials.
% correct	Total number of correct matches/ total number of trials.
Perseverative responses	The number of incorrect responses that would have been correct for the preceding category/rule.
Perseverative errors	The number of errors in which the participant has used the same rule for their choice as their previous choice.
% perseverative errors	Perseverative errors/ total number of trials.
Non-perseverative errors	After excluding the perseverative errors, the number of other errors.
Total errors	sum of perseverative + nonperseverative errors.
Percent of total errors	total errors / total number of trials.
# trials to complete 1st rule	Total number of trials needed to achieve the first 10 consecutive correct choices.
# rules achieved	The number of runs of 10 consecutive correct choices.
Failure to maintain set	The number of times 5 or more consecutive correct choices occur without completing the category (ie, without reaching 10 consecutive correct choices).

Table 2.1: WCST variables and their definitions (Foundation, 2008).

WCST performance is correlated with real-world decision performance (Davis et al., 2011). Reaction time on the complex decision scenarios was correlated with failure to maintain set, whereas inconsistency was negatively correlated with percent correct and positively correlated with percent errors and non-perseverative errors. These findings suggest that our wargaming version of the WCST should predict which soldiers demonstrate optimal decision-making during more complex wargaming scenarios, as well as the types of errors made that

impeded optimal decision-making.

Importantly, trial by trial EEG activity can be used to detect the transition from exploration to exploitation in the WCST. When EEG activity during early trials is compared to late trials within a set of trials with the same matching rule, distinct early vs late patterns of frontal and nonfrontal lobe activity (Barceló and Rubia, 1998). Late trials are characterized by a large Pb3 wave in the mid-parietal areas. Furthermore, EEG measurements from healthy participants can distinguish between perseverative errors and nonperseverative errors and between distractions and nondistractions where behavioral data (ie, reaction time) can not (Barceló, 1999). Perseverative errors showed an absence of N1, reduced P2 waves, and larger P3b waves than did non-perseverative errors. Distractions showed significantly larger P2 waves than non-distractions. Notably, perseverative errors and distractions EEG patterns differed from each other, with perseverative errors located in the posterior and distractions located in the frontal central brain areas. In sum, the IGT and the WCST each have components that enable us to address study 1 questions listed in the introduction. In the next section, we outline several possible statistical models for characterizing the mechanics of optimal decision making.

2.4. Mechanics of Optimal decision-making

Optimal decision-making in a human decision-maker can be described as the ability to identify the action in a given situation, or state of the world, that maximizes the long-term expected utility of the individual (or group) and the mechanics of optimal decision-making can be thought of as the manner is which the human uses his internal information processing system to perceive, reason about, and arrive at a decision. Reinforcement learning provides an empirically inspired conceptual model of how biological organisms learn, through trial and error, the appropriate action to take in a given state in order to maximize long-term reward. (Sutton and Barto, 1998; Thorndike, 1911; Skinner, 1938). Placed in the context of an information processing view of cognition, reinforcement learning would provide one way of representing manner in which experience is accrued over time.

2.4.1. Measure of Performance: Regret

A main goal of study 1 is to define an objective measure of performance for subjects participating in the military version of the IGT that goes beyond typical IGT decision performance measures, such as advantageous selection bias. The designers of the test know the pay-out schedule of each deck in advance, so it is possible to know at any point in the sequence of trials which deck provides the best reward, which is unknown to the players. Using this information it should be possible to calculate the difference between the reward achieved by a subject at a given trial and the best possible reward available at that turn. This difference is referred to as regret.

Regret is often used as a performance metric for multi-armed bandit problems (Szepesvari,

2010; Sutton and Barto, 1998). The goal in a multi-armed bandit problem is to maximize the total payoff obtained in a sequence of allocations (Lai and Robbins, 1985). The problem is often described as a sequential allocation problem, sequential sampling problem, or sequential decision-making problem and was inspired by the problem of a gambler facing a collection of slot machines, each with a different and initially unknown probability of winning along with an equally unknown payout or reward (Auer et al., 2002; Szepesvari, 2010; Auer and Ortner, 2010). The player only receives information regarding each arm by playing the arm and collecting an observation. This forces the player to balance the desire to find the arm with the greatest long-term value with the need to get acceptable immediate rewards on each play. This need to balance exploration and exploitation in sequential experiments has attracted interest in bandit algorithms from a variety of fields (clinical trials, online advertising, and most recently information packet routing) (Auer et al., 2002).

Regret is the expected loss after n played due to the fact that a player does not always choose the best route. In order to compare the performance of players and algorithms designed to solve bandit problems, the difference between the total reward received from playing the best arm available at each sequential trial and the actual reward accrued by the player is determined, this can be done over any horizon n of trails. Lower regret is better.

2.4.1.1. Method 1: Absolute Regret

This calculation compares the outcome of player actions to the outcome generated by playing the optimal policy at each of the n trials. Given $K \geq 2$ routes and sequences $r_{i,1}, r_{i,2}...$ of unknown rewards associated with each route $i = 1, ..., K$, at each trial, $t = 1, ..., n$, players select a route I_t and receive the associated reward $r_{I_t,t}$. Let $r^*_{i,t}$ be the best reward possible from route i on trial t ; (Auer and Ortner, 2010). The regret after n plays $I_1, ..., I_n$ is defined by

$$R_n = \sum_{t=1}^{n} r^*_{i,t} - \sum_{t=1}^{n} r_{I_t,t}. \tag{2.1}$$

2.4.1.2. Method 2: Pseudo Regret

Pseudo regret is used in analysis of stochastic bandit problems. Pseudo regret makes use of expectations. For $i = 1, ...K$ we denote μ_i as the mean reward of route i and let $\mu^* = \max_{i=1,...,K} \mu_i$. Then the pseudo regret is determined by examining the difference in expected value of the best arm and the expected value of the arm chosen (Audibert and Bubeck, 2010).

$$\overline{R_n} = n\mu^* - \sum_{t=1}^{n} \mu_{I_t} \tag{2.2}$$

Regret per Turn When used as a performance metric for bandits regret provides insights when used in the aggregate over the course of a set of n trials, total regret, and when examined per turn. Thus, as a player identifies routes with better returns the regret accumulated per turn should go down. Regret per turn in particular can provide a measure of how well a player is

balancing exploration and exploitation and a measure of the player's ability to identify the best arm available at a given point in time.

2.4.2. Reinforcement Learning

The term reinforcement learning (RL) in this context refers to a subfield of machine learning that relies on trial and error learning to determine the best action, $a \in A$, to choose in a given state, $s \in S$, in order to maximize the sum of a numeric reward signal provided by the environment in which it is operating. Reinforcement learning systems were inspired by research in animal behavior and are convenient choices to serve as simulated human players. RL systems usually include a reward function that maps states to rewards, a value function used to determine the long-term value of a state or state-action pair, and a policy to guide action selection while balancing exploration and exploitation (Sutton and Barto, 1998; Russell and Norvig, 2010). For our purposes, with our version of the IGT, we can employ one of the action-value estimation methods to simulate the manner in which a human keeps a running estimate of the value of each alternative in conjunction with one of the exploration policies. Note that most of action value estimation methods employ some discount parameter, allowing us to account for the recency bias (recalling and choosing the most event/selection) in human decision-making and that both of the most commonly used exploration policies include a parameter balancing the ratio of exploration and exploitation. Post experiment it will be interesting to fit the parameters of the RL algorithms to the data from the experiments to represent human decision-making in this task (Walsh and Anderson, 2010).

2.4.2.1. Exploration policies

Two of the more commonly used policies to balance exploration and exploitation are the ϵ-greedy method and the softmax method (alternatively referred to as Boltzmann exploration). Note both of these are considered stochastic exploration policies (Sutton and Barto, 1998).

ϵ-greedy One of the alternatives to a pure greedy policy is ϵ-greedy, where an exploration rate specifying the probability of selecting a non-greedy action, $\epsilon \in (0, 1]$, is specified in advance. Setting ϵ is fairly intuitive since it can be thought of as the fraction of time that the policy will choose a non-optimal action.

Boltzmann Exploration Rather than simply choose an action at random, Boltzmann exploration makes use of the estimated value of the action, $Q(s, a)$, where s is the state and a is the action, making the probability of choosing an action proportional to its estimated value.

$$P(a_i|s) = \frac{\exp^{\frac{Q(s,a_i)}{\tau}}}{\sum_i \exp^{\frac{Q(s,a_i)}{\tau}}} \tag{2.3}$$

The variable τ serves as a scaling parameter making the probability of selecting a greedy action go toward 1 as $\tau \to 0$ and producing a more exploratory sampling policy with larger values of τ. Identifying good values for τ can be more difficult than ϵ.

2.4.2.2. Action value estimation

Two popular methods for estimating the value of a state-action pair, $Q(s, a)$, are Q-learning and State-action-reward-state-action (SARSA) (Sutton and Barto, 1998; Bertsekas and Tsitsiklis, 1996). Both come with theoretical guarantees of convergence to the optimal policy given that they are paired with an exploration policy that guarantees all state-action pairs will be visited infinitely often - which for most practical applications is not particularly useful. Q-learning and SARSA have been used by previous researchers to model human performance in multi-arm bandit settings (Walsh and Anderson, 2010). In our task we have a single state with four possible actions.

Q-learning Q-learning is a model free temporal differencing method of estimating the value of a state-action pair (Watkins, 1989). Let,

$$\delta = r_{t+1} + \gamma max_a Q(s_{t+1}, a_t) - Q(s_t, a_t) \tag{2.4}$$

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha\delta \tag{2.5}$$

, where $\gamma \in (0, 1]$ is a discount factor and $\alpha \in (0, 1)$ is referred to as a learning rate or step size parameter.

State-action-reward-state-action (SARSA) SARSA only differs from Q-learning in the update to the δ term (Sutton and Barto, 1998). Let,

$$\delta = r_{t+1} + \gamma Q(s_{t+1}, a_{t+1}) - \ Q(s_t, a_t) \tag{2.6}$$

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha\delta, \tag{2.7}$$

, where $\gamma \in (0, 1]$ is a discount factor and $\alpha \in (0, 1)$ is referred to as a learning rate or step size parameter.

Note: determining *good* values for α and γ for either algorithm takes some amount of thought or experimentation and varies based on the environment. For our purposes we could choose either Q-learning or SARSA (or both).

Upper Control Bound (UCB) There are several variants of the UCB algorithm, we'll choose UCB1(Auer et al., 2002). The upper control bound (UCB1) algorithm uses a deterministic exploration policy rather than a stochastic one (ϵ-greedy or Boltzmann). The UCB1 algorithm always chooses the route that maximizes

$$\overline{r_i} + \sqrt{\frac{2 \ln n}{n_i}} \tag{2.8}$$

where $\overline{r_i}$ is the average reward obtained from route i, n_j is the number of times route j has been played so far, and n is the overall number of plays done so far.

2.5. Sequential Detection Methods for Detecting Exploration-Exploitation Mode Changes

Sequential Detection Methods can be used to identify when a subject switches from *exploration* to *exploitation* mode or vice-versa. To make such a determination, we have three types of data available:

- The routes subjects select,
- The time between clicks (i.e., the *latency*), and
- The outcome of each selection (i.e., damage to friendly forces).

We initially focused on selection outcomes. At this point we think the two most relevant measures are subjects' choices of routes and time between clicks, which we refer to as trial latency.

With the choice of routes, the notion is that during exploration the subject will be trying different routes to gather data about the best route and over time they will eliminate routes until they only use one. Once they determine a particular route is best they then enter into exploitation mode by continuously choosing that route. However, they may go back into exploration mode if they find out that they made a poor choice of route to exploit, subsequently pick and focus on a new route, and then go back into exploitation mode. Thus, what we expect to see during exploration is the alternating selection of various routes, while during exploitation we expect the subject to focus exclusively, or nearly exclusively, on one route.

In terms of time between clicks, we expect that during exploration the subject is purposely choosing routes and actively keeping track of the outcomes on the routes, while during exploitation the subject no longer has to spend as much time thinking about the other routes. Thus, we expect that latency will be greater during exploration than during exploitation. However, as with the choice of routes, should it turn out during a period of exploitation that the route choice is suboptimal, then the subject may revert to exploration mode and we would expect to see an increase in latency as he or she goes back to thinking about and collecting information on the other routes.

What both of these variables have in common is that what we're looking for is a decrease in variability, either in terms or routes or latency, as a subject goes from exploration to exploitation (and an increase if there is a reversion from exploitation to exploration).

2.5.1. Defining Some Sequential Methods

Let's start with defining some methods for latency, which should be conceptually a bit more straightforward, and then move on to some methods for route selection.

2.5.1.1. Latency

The idea here is that as a subject moves from exploration to exploitation the time between clicks should decrease, as should the variation in time between clicks. This suggests that monitoring either the mean time between clicks or the click variance (or both) could be informative.

2.5.1.2. Method 1: The Exponentially Weighted Moving Average

Let's start with the former, where we could use the *exponentially weighted moving average* (EWMA) method drawn from the statistical process control literature (Fricker, 2010). Let x_i denote the latency at time i, $i = 2, 3, \ldots, 100$ (where, presumably, there is no latency at time $i = 1$). Then at time i we would monitor

$$E_i = \alpha x_i + (1 - \alpha) E_{i-1},$$

where α is a smoothing parameter, $0 < \alpha \leq 1$ and typically the method starts by setting $E_1 = x_2$. Here we assume that at time $i = 1$ the subject starts out in the exploration mode and the question is to identify when he or she switches to exploitation. This is done by setting a threshold h and the first time i that $E_i < h$ we declare that the subject is now in exploitation mode.

Three questions then arise: (1) How to choose α? (2) How to choose h? and (3) Is h subject specific? The last one is particularly important because if each subject has a different threshold then it's likely that we can't use this methodology. The pilot data will assist in answering to these questions. If it seems that subjects are largely similar, we might be able to use the pilot data along with some simulation to pick h. Similarly, we can use some known results and simulation to help us pick an appropriate α. If h is relatively stable across subjects, then we also can determine how to identify a subject that switches back from exploitation to exploration.

2.5.1.3. Method 2: Monitoring Sequential Sample Variances

Method 1 leads us to the idea of monitoring latency variance which may be easier to implement than monitoring the mean since, when a subject goes into exploitation mode, it is possible that the variance will get close to zero (for all subjects). This method is one way to implement a sequential scheme where we would monitor the sample variance calculated from moving windows of data. Specifically, as before let x_i denote the latency at time i, $i = 2, 3, \ldots, 100$. Then for some window of data of size $w + 1$, starting at time $i = w + 2$, sequentially calculate

$$s_i^2 = \frac{1}{w} \sum_{j=i-w}^{i} (x_j - \bar{x}_i)^2,$$

where

$$\bar{x}_i = \frac{1}{w+1} \sum_{j=i-w}^{i} x_j.$$

The idea is to monitor $s^2_{w+2}, s^2_{w+3}, s^2_{w+4}, \ldots$ and when it is less than some threshold h we declare that the subject has gone from exploration to exploitation.

For this method, the question is how to choose w. There are two considerations: (1) $w+1$ should be smaller than the smallest length of time a subject might be in exploration mode when the experiment first starts, and (2) smaller is better in the sense that the method will more quickly indicate the shift to exploitation, but $w+1$ cannot be so small that the sample standard deviation estimates are too variable because of excess noise. Ultimately, we will want to do some simulations to see what a good choice for w might be. Our initial guess would be something in the range $4 \leq w \leq 8$ or so.

Now, there is also the question of how to detect whether someone reverts from exploitation back to exploration. One possibility would be to continue to monitor the sample variances and, once someone is in exploration mode, should $s^2_i > h$ then we say they have reverted back to exploration. However, it may be that we need two thresholds, call them h_1 and h_2, where $h_2 > h_1$, that would work as follows. For someone in exploration mode, then they only switch to exploitation at time i when $s^2_i < h_1$ while for someone in exploitation mode, they only switch to exploration at time i when $s^2_i > h_2$. The key idea here is that having two thresholds with some separation between them may decrease inadvertent (i.e., excessive) switching back and forth between modes due to noise in the data.

2.5.1.4. Route Selection

In a sense, the idea with monitoring route selection is similar to that of latency. We're looking for when a subject stops switching between routes and starts to concentrate on a single route. A simple retrospective way to do this would be to specify a rule that says sequences of matching route selections greater than some number m represent exploitation (and otherwise the subject is in exploration mode). This method could be a good "baseline" method against which to evaluate all others.

2.5.1.5. Method 3: Monitoring the Rate of Route Switches

The idea with Method 3 is the rate at which a subject is switching between routes could be an indicator of exploration mode. Let r_i denote the route chosen by a subject at time i, $i = 1, 2, \ldots, 100$, where r_i can take on values $1, 2, 3$ or 4. Then let y_i be an indicator variable denoting that the route at time i does not match the route chosen at time $i-1$. That is, $y_i = 0$ if $r_i - r_{i-1} = 0$ and otherwise $y_i = 1$.

Then, as with Method 2, for some window of data of size $w+1$, starting at time $i = w+2$, sequentially calculate

$$\bar{y}_i = \frac{1}{w} \sum_{j=i-w}^{i} y_j.$$

Then we monitor $\bar{y}_{w+2}, \bar{y}_{w+3}, \bar{y}_{w+4}, \ldots$ and when the rate is less than some threshold h we declare that the subject has gone from exploration to exploitation.

Similar to Method 2: (1) $w+1$ should be smaller than the smallest length of time a subject might be in exploration mode when the experiment first starts, and (2) smaller is better in the sense that the method will more quickly indicate the shift to exploitation, but $w+1$ cannot be so small that the sample means are too variable because of excess noise. Compared to Method 2, Method 3 has a chance of performing better simply because means are more accurately estimated with small samples compared to standard deviations.

2.5.1.6. Method 4: Monitoring the Variance of the Rate of Route Switches

This approach is similar to Method 2 but applied to the variance of $z_i = \sum_{j=i-w}^{i} y_j$, the number of times a subject switched routes in the past $w+1$ time periods. The idea is that during exploration the variance of z_i would be high because the subject is alternately switching between routes and sometimes trying a route for a couple of times in a row. On the other hand, the variance will drop off once exploitation starts because switching will stop. (Note that a potential weakness of this method is that the variance of z_i could also be low during times when the subject continues to switch back and forth between routes consistently.)

2.5.1.7. Method 5: Monitoring the Number of Unique Routes Tried

The rationale behind Method 5 is a subject who is selecting between more routes is more likely to be in exploration mode. As before, let $w+1$ be a window of historical data and let n_i be the number of unique routes chosen at times $i, i-1, \ldots, i-w$. So n_i can take on values $1, 2, 3, 4$. Then we monitor $n_{w+2}, n_{w+3}, n_{w+4}, \ldots$ and when the rate is less than some threshold h we declare that the subject has gone from exploration to exploitation. In this case, it could be that we set $w = 2$ or 3 and exploration occurs when $n_i = 1$.

The same questions arise here as to the appropriate choice for $w+1$ and the thresholds, as well as whether there should be two thresholds, one for exploration to exploitation and another for exploitation to exploration.

2.5.1.8. Discussion

Note that all five of the methods above are prospective methods, meaning at time i they only use information from (at most) times $1, 2, \ldots, i$. That is, they do not use information from the "future." However, given that we will be retrospectively evaluating our subjects' data, we could use "future" data such as in the the simple retrospective method first described in Section 2.5.1.4. We should think about whether we want to also look at these types of retrospective methods. The benefit of the prospective methods is that they could be applied in real time to subjects as they're engaged in the experiment to (help) determine what mode they are in. In sum, there are a plethora of ways in which to define the transition from exploration and exploitation. Through pilot data collection, we will be able to implement and compare the most appropriate methods.

3. STUDY 1 METHODS

3.1. Experimental Design

3.1.1. Population of Interest

Our population of interest is active duty military personnel. Recruitment will occur through bulk email to all NPS students, faculty and staff, posting of flyers and word of mouth. Additionally, the PI will briefly present the study to various campus groups. Figure B.1 contains an example of a recruitment flyer.

3.1.2. Decision-making Tests

For study 1, the decision-making tests are the convoy task and map task.

3.1.2.1. Convoy Task

Our version of the IGT, the convoy task, serves as a simple wargame for this project. In the convoy task, subjects are asked to select one of four possible routes over an unknown number of trials to maximize the damage to enemy forces while minimizing the friendly damage accrued over all trials. These routes are analogous to the decks of the original IGT.

At each trial, the subject is provided immediate feedback in the form of three separate pieces of information: a reward, penalty and a running total. The reward, number of enemy damage, is called *Damage to Enemy Forces*. The penalty, the number of friendly damage, is called *Damage to Friendly Forces*. The running total is called *Accumulated Damage*, defined as the previous trial's value of Accumulated Damage plus the previous trial's Damage to Enemy Forces minus the previous trial's Damage to Friendly Forces. The units of value are in damage. Damager To Enemy Forces is considered positive in value (damage given to the enemy) and desirable to the player. Damage to Friendly Forces is negative in value (value lost due to damage to friendly forces) and is not desired by the player.

The feedback for the convoy test is derived from the first published IGT. The convoy task payout schedule for each route demonstrated in Table D.10 is constructed from the original IGT schedule demonstrated in Figure 2.6. With the few pilot trials to go, we should observe if the players notice the pattern. Each route has its own 'deck', a scripted, ordered set of specified values. For example, every player will find that the third time they pick deck A, it returns +100 and -150. Even though these returns by deck are set and the same for each player, the games will progress differently due to the divergence of deck selection between players.

The convoy task offers minimal visual difference between images representing the available options. The intent of similar looking options is to minimize the visual bias. This is consistent with the first IGT by Bechara et al. (2005) in Figure 2.7.

Route A		Route B		Route C		Route D	
Min.	-250	Min.	-1150	Min.	0	Min.	-200
1st Qu.	-150	1st Qu.	100	1st Qu.	0	1st Qu.	50
Median	25	Median	100	Median	25	Median	50
Mean	-25	Mean	-25	Mean	25	Mean	25
3rd Qu.	100	3rd Qu.	100	3rd Qu.	50	3rd Qu.	50
Max.	100	Max.	100	Max.	50	Max.	50

Table 3.1: Summary statistics for the convoy task.

Figure 3.1: Screen shot of the convoy task in piloting, a typical subject's view of the task. We see the player's last choice caused 100 damage to the enemy (*Damager To Enemy Forces*) and a loss of -250 to friendly forces (*Damage to Friendly Forces*) resulting in a trial loss of -150 (not shown). The *Accumulated Damage* is 2750. A positive *Accumulated Damage* value is desirable to the player. Notice four routes are represented by the same image.

The subject seeks to determine which route to select on the next turn through repeated sampling of routes. A player selects routes until the end, unknowing it will complete after 200 selections. The assumption is that the subject maintains some estimate of the value similar to Accumulated Damage for each route and updates the estimate after each trial. The accuracy of the estimate will vary between subjects as will the manner in which the subjects incorporates information indexed by trial into their estimate.

3.1.2.2. Map Task

Our version of the WCST, the map task, serves as a simple wargame for this project. In the map task, subjects view 5 maps, one map displayed at the top center of the screen, the remaining four displayed across the bottom of the screen. Figure 3.3 is a typical subject's view of the task. The maps are analogous to the cards of the original WCST. Each map contains military graphic control graphics that vary in meaning, color and shape. These graphics are described in Figure 3.2, and developed from FM 1-02, Operational Terms and Graphics. Subjects are asked to match one of four lower maps to the top one over an unknown number of trials.

	friendly graphics	intent graphics	enemy graphics
Level 0	no friendly graphic	no intent graphic	no enemy graphic
Level 1	 friendly armor platoon	 ambush	 enemy infantry squad
Level 2	 friendly aerial vehicle	 clear	 enemy anti-armor squad
Level 3	 friendly infantry platoon	 block	 enemy anti-air squad

Figure 3.2: Description of graphics in map task. There are three categories of graphics, friendly (colored blue), intent (colored black) and enemy (colored red). The sorting rules correspond to the same categories. Each category has four levels, each with a particular corresponding graphic.

Over several trials, participants try to figure out the matching rule that will correctly match the map on the top of the screen with one of the four maps at the bottom of the screen. This process of match maps is similar to card matching in the original WCST. Unbeknownst to the participant, the matching rule changes once the participant has 10 consecutive correct matches. For example, after 10 consecutive correct matches sorting the maps using the sorting rule based on the friendly graphic, the matching rule changes to sorting maps according to the intent graphic. The task is completed when either the participant has successfully completed two rounds of each matching rule or until they have completed 128 trials. For the map task, we use the same decision performance measures developed from WCST, described

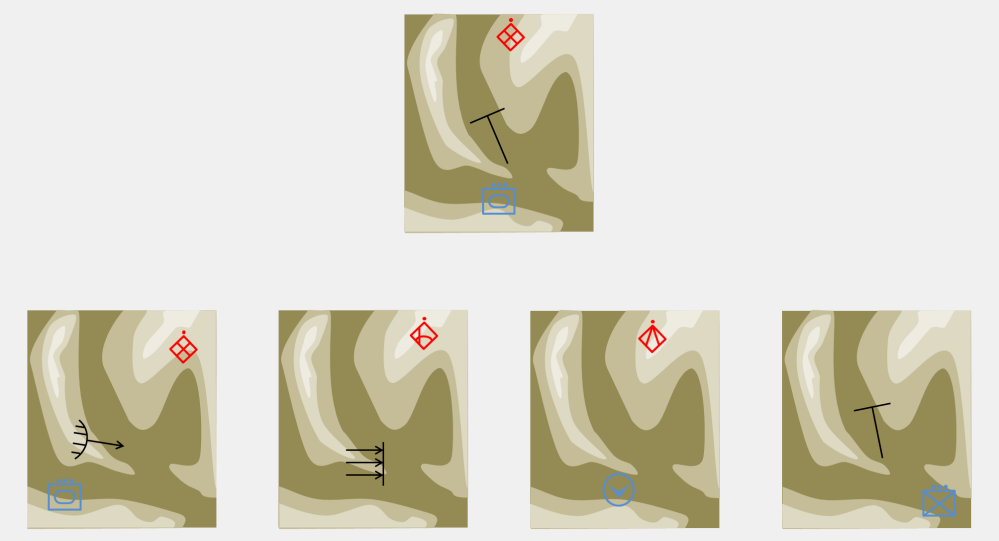


Figure 3.3: Screen shot of the map task in piloting, a typical subject's view of the task.

in Table 2.1.

In previous work described in section 2.3.2, reaction time on the complex decision scenarios was correlated with failure to maintain set, whereas inconsistency was negatively correlated with percent correct and positively correlated with percent errors and non-perseverative errors (Davis et al., 2011). These findings suggest that our wargaming version of the WCST should predict which soldiers demonstrate optimal decision-making during more complex wargaming scenarios, as well as the types of errors made that impeded optimal decision-making.

3.1.3. Surveys

We use surveys to quantify and categorize blocking factors such as elements of military experience and to collect qualitative responses from the subjects at the conclusion of the tasks. We use two surveys to accomplish this, a demographic survey at the beginning of the experiment, and an post task survey at the end.

3.1.3.1. Demographic Survey

The demographic survey in Figure D.2 is administered prior to the decision-making tasks. The survey includes questions regarding participants deployment history, as well as general demographic information such as age and rank. Through this survey, we will double check that the participants are indeed active duty military as well as check that results aren't due to certain subject demographic characteristics.

3.1.3.2. Post Task Survey

The post-task survey in Figure D.8 is administered after the completion of the decision-making tasks. Participants provide qualitative responses regarding their strategies for each decision-making task.

3.1.4. Covariate Measures

Because the decision-making tasks place demands on working memory and visual processing speed, we are including covariate measures of these cognitive functions. The tasks also are highly visual; therefore, a visual acuity test also is administered.

3.1.4.1. Digit Span Memory Test

Digit span forwards and backwards test Bechara et al. (1994) measures working memory. In digit span forwards, the experimenter states a series of digits, starting with 2 digits, and the participant must repeat them back. The number of digits increases, with two trials per number of digits. The test is discontinued if the participant has an incorrect response to both trials for a particular number of digits. In digit span backwards, the same procedure is followed, except this time the participant must repeat the digits in the reverse order. Figure D.3 contains the instructions for digit span forward and backward.

3.1.4.2. Trails A, Trails B

Trails A and B test visual processing speed (Grant and Berg, 1948). In Trails A, the numbers 1 through 25 are randomly distributed on the paper as demonstrated in Figure D.6. The participant starts at 1 and must draw a line to each number in chronological order. Participants are instructed to work as quickly and accurately as they can. In Trails B, participants now see both numbers and letters and must connect 1 to A, A to 2, 2 to B and so on until they reach Z as demonstrated in Figure D.7. They also are instructed to work as quickly and accurately as they can.

3.1.4.3. Snellen Test

Because the decision tasks are visually based, the Snellen eye chart is used to measure subjects' visual acuity at the beginning of the experiment. The Snellen eye chart D.3 is placed on the wall and consists of 11 lines of block letters, in which each line of letters gets progressively smaller. Subjects stand 20 feet from the chart, cover one eye, read aloud as many lines as they can. They then cover the other eye and read aloud as many lines as they can. The experimenter records the last line that the subject could accurately read for each eye.

3.1.5. Equipment

The devices used in this study consist of a laptop computer, two eye tracking stereo cameras, a desktop computer, and an electroencephalogram (EEG). The laptop runs FaceLAB 5.0.7

software on a Windows XP operating system. The stereo cameras supply data to FaceLAB on the laptop. FaceLAB software and the stereo cameras were made by Seeing Machines Inc. The desktop computer runs the EyeWorks data collection suite and Advanced Brain Monitoring (ABM) Visual software on the Windows 7 operating system. The laptop has a 15" screen that is not viewed by the subjects. The desktop uses a 30" primary monitor which is viewed by the subjects, and a 24" secondary monitor which is not viewed by the subjects.

The stereo cameras use 12 mm lenses to detect infrared light reflected off the subjects' eyes and face to monitor the position of the head and direction of the eye gaze. This data is fed from the laptop to the EyeWorks Record software on the desktop.

EEG data is recorded through an ABM X10 B-Alert Headset through 9 channels (F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4) and sent through wireless connection to B-Alert Visual software on the desktop.

Other materials used include 70% ethyl alcohol to clean the subjects' mastoid reference points, Synapse brand electrolytic gel, and recording electrodes provided by ABM.

3.1.6. Procedures

The subjects, all volunteers, complete the experiment in a single visit. Upon arriving to the test location, they complete a demographic survey, consent to participate form (found in Appendix D.1) and the baseline and cognitive tasks including the digit span forward/backward task, and two forms of the trail making test (TMT). Next they calibrate the EEG and eye-tracking systems. Eye tracking calibration includes verifying the integrity of the camera configuration, building of a personalized head model for the subject, and calibrating the subject's gaze with respect to the screen. EEG calibrating tasks include getting scalp and reference impedance levels under 40 kOhms and creating a baseline EEG profile using the 3-choice vigilance, eyes open, and eyes closed tasks. Once all calibration steps are satisfied, the subject completes the convoy task and the map task. With the tasks complete, they complete the post task survey and are reminded of the confidential nature of the data collected. Full procedure notes are found in Appendix D.2.

3.2. Analytic Methods

Figure 3.4 demonstrates the proposed methodology of analyzing the data. The research team will separate into two analysis teams, the neurophysiology and decision analytics team. The neurophysiology team will focus on leveraging psychology and human factor techniques to answer research questions listed in the introductory section. The decision analytics team will focus on leveraging machine learning, process control and simulation techniques to answer the research questions. The two teams will compare and combine their procedures in developing a comprehensive approach to identifying transitions between exploring and exploiting

information.

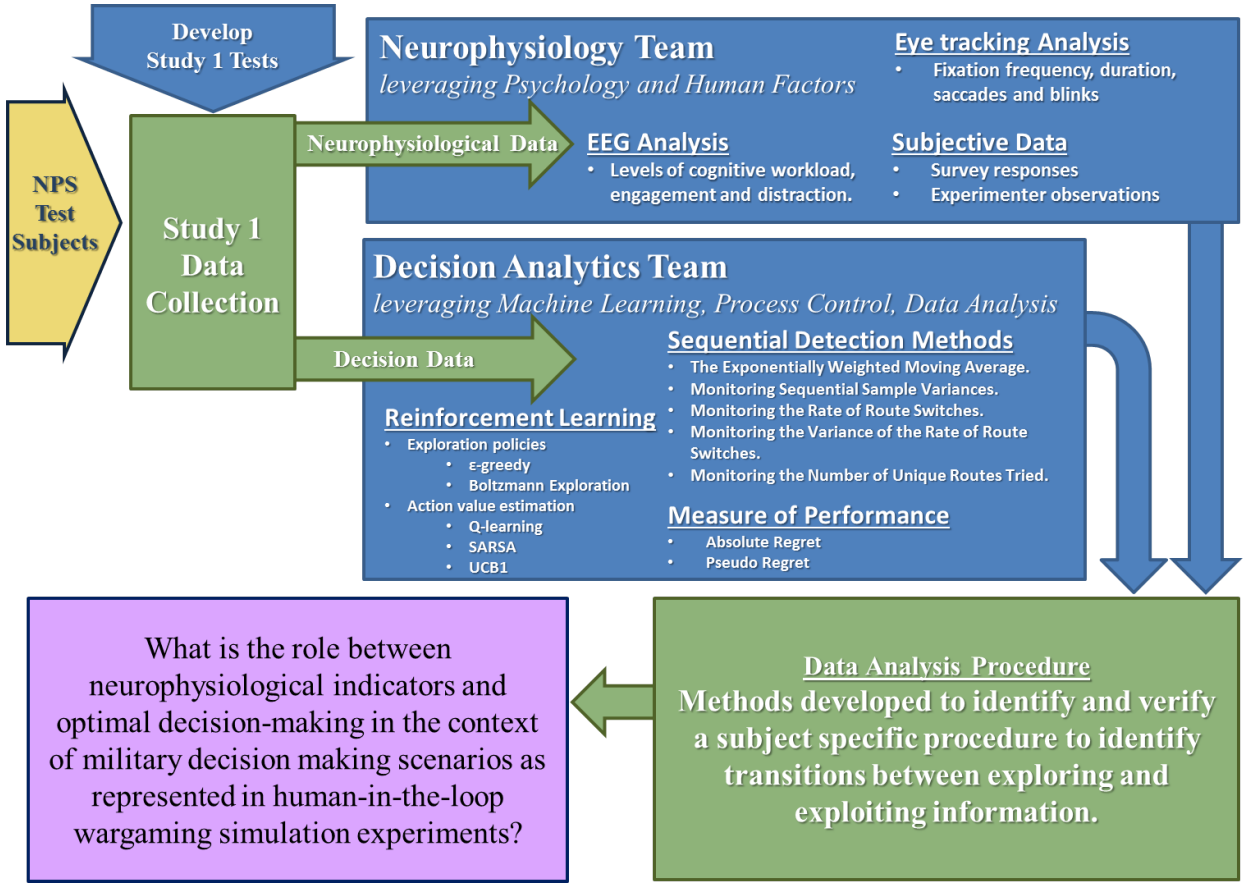


Figure 3.4: Data analysis methodology. The research team will separate into two analysis teams, neurophysiology team and decision analytics team combining their results to verify and support each other with applicable indicators.

4. STUDY 1 PILOT DATA COLLECTION

The purpose of the pilot study was to confirm the validity of the convoy and map tasks, as well as to ensure that the tasks did not have ceiling or floor effects. Towards achieving validity, several procedures were conducted. First, subject matter experts (SMEs) provided feedback regarding the extent to which the two tasks are adapted to include a military scenario. Second, the SMEs confirmed that the tasks tapped reinforcement learning and cognitive flexibility. Third, the convoy task was designed so that the distributions of enemy damage and friendly damage per route selection in the convoy task was the same as the distribution of gains and losses per deck in the IGT. Similarly, the number of icons on which to match, the number of consecutive correct selections before the sorting rule changes, and the number of categories needed to complete the map task is the same as in the WCST. Fourth, the same general instructions are used in our tasks as in the original tasks. We also have created decision performance measures that are operationally identical to those typically used in the IGT and WCST. These procedures were conducted over numerous iterations.

Finally, we conducted a small pilot study with the finalized versions of the convoy and map tasks. Preliminary results from pilot data on the decision performance measures are consistent with results based on the IGT and WCST and demonstrate reasonable amounts of variability on these measures. Below, we describe these preliminary results for each task.

4.1. Verification of Convoy and Map tasks

4.1.1. Verification of Convoy Task

We explored several different possible measures of overall decision-making performance, such as final damage score, frequency of damage, advantageous selection bias, and trial latency. Table 4.1 below depicts how well each pilot subject performed on these measures.

Final Damage All subjects start with 2000 enemy damage. Therefore, the Final Damage is calculated as the difference between the initial Damage Score and the last Damage Score. A difference greater than 0 demonstrates optimal decision performance, whereas negative scores indicate suboptimal decision performance. Pilots' Final Damage scores ranged from -100 to 1550. Thus, pilot 3 showed superior decision-making, pilot 1 showed optimal decision-making, and pilots 2 and 4 demonstrated suboptimal decision-making.

Frequency of Damage Frequency of damage is defined as the number of trials in which friendly damage occurred. In comparing frequency of damage with Final Damage score, it is seen that relatively low frequency of damage does not necessarily correlate with a good Final Damage score. To examine this further, we also noted recorded the frequency of trials with heavy friendly damage (1250 damage). Of note, pilot 3, who had the best Final Damage,

Performance Variables	Pilot 1	Pilot 2	Pilot 3 (1st 100 trials)	Pilot 4 (1st 100 trials)
Final Damage	350	-100	1550	-50
# trials with friendly damage	23	13	26	26
# trials with heavy friendly damage	4	5	2	4
Route selection frequency				
Route 1	11	7	10	11
Route 2	43	50	27	42
Route 3	26	3	30	27
Route 4	20	40	33	20
Advantageous selection bias	-8	-14	26	-6
mean latency per trial (sec) (SE)	2.039 (1.037)	2.083 (1.466)	4.110 (1.722)	3.668 (2.354)
median latency per trial (sec)	0.768	0.312	1.577	1.123
mode latency per trial (sec)	0.668	0.250	1.025	0.655

Table 4.1: Descriptive statistics of pilot subjects performance on convoy task. Note, pilot subjects 1 and 2 completed the initial 100 trials whereas pilot subjects 3 and 4 completed 200 trials.

also had the lowest proportion of friendly damage trials that incurred heavy damage.

Advantageous selection bias The typical decision performance measure from the IGT is the advantageous selection bias, in which the proportion of bad routes selected is subtracted from the proportion of good roads selected. According to the IGT, routes 3 and 4 are considered good; 1 and 2 are considered bad. Positive advantageous selection bias scores indicate a propensity to select the good routes, whereas negative scores indicate a tendency to select the bad routes. According to this measure of decision performance we find that only pilot subject 3 shows optimal decision performance. We note that although pilot subject 1 had an optimal Final Damage score, their advantageous selection bias score was suboptimal.

Route selection Route selection is the frequency with which the subject selected each route over all trials. Results are consistent with previous work by Steingroever et al. (2013), in which high individual variability in frequency of selected routes was found. The one exception is that all pilot subjects avoided route 1. In Figures 4.1 and Figure 4.2, we illustrate the variability in route selection and trial latency between pilot subjects. Pilot subject 3 frequently switched routes, yet clearly favored routes 3 and 4, possibly leading to fewer number of trials with heavy friendly damage.

Trial Latency Latency is defined as the amount of time subjects take to make a decision on each trial. It is measured as the amount of time taken between key press selections from trial to trial. Mean latencies are higher than median latencies because all pilots took at least 100 seconds to make a decision on the first trial. Thus, the median and mode latencies more accurately reflect pilots latencies over the course of the 100 trials (part of that time is due to task instruction). As would be expected, individual differences in trial latency is evident.

Importantly, no subject had a trial latency below 150 milliseconds, a typical minimum cutoff for intra-individual variability in reaction time to exclude possible measurement error due to accidental key press or distraction (Bielak et al., 2010).

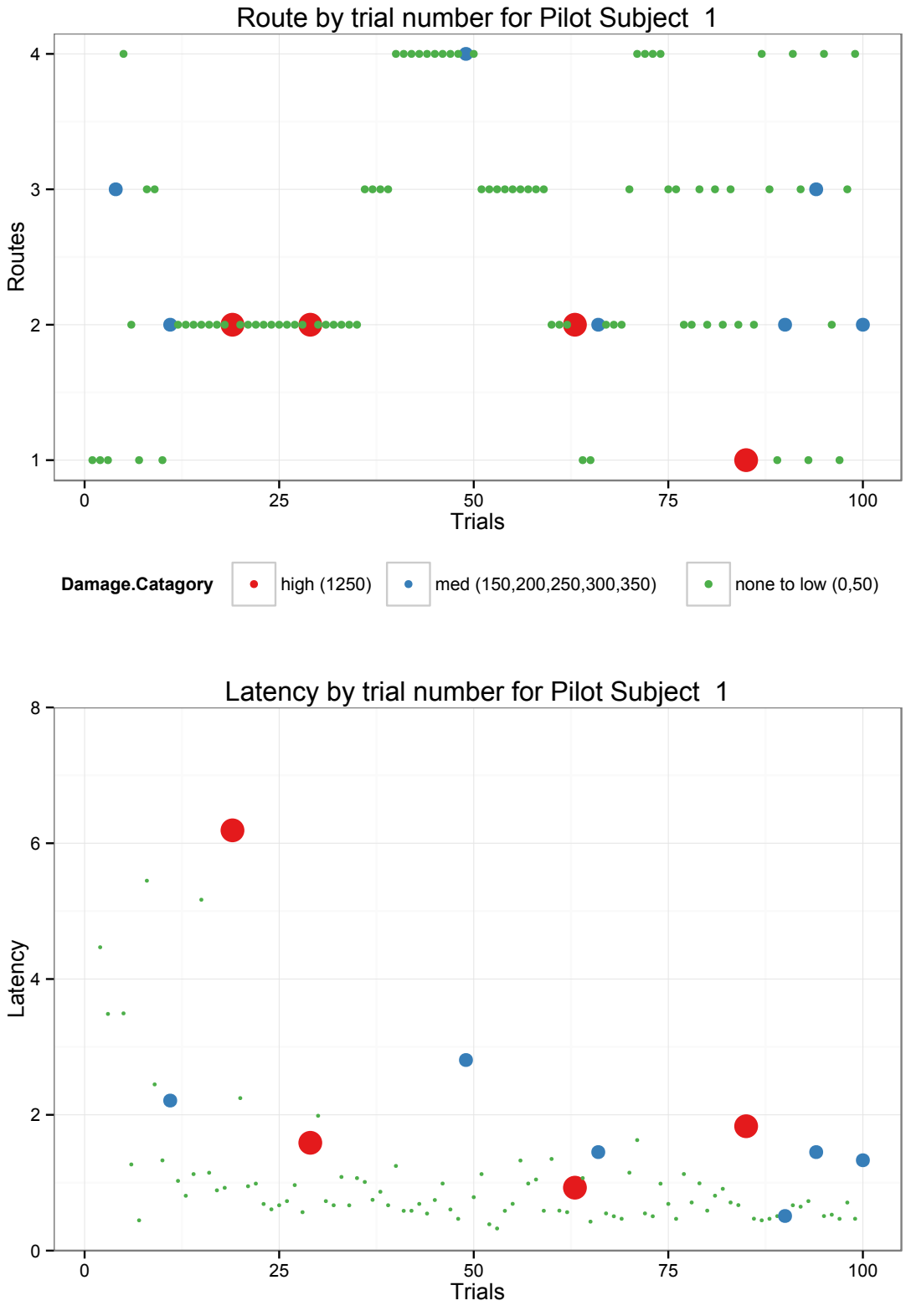


Figure 4.1: Route selection by trial for pilot subject 1 (top). Latency time by trial for pilot subject 1 (bottom). Damage.Category is the level of net damage received (Damage to Friendly Forces - Damage to Enemy Forces) on the previous trial. For example, the large red circles represent the next decision after receiving high friendly damage.

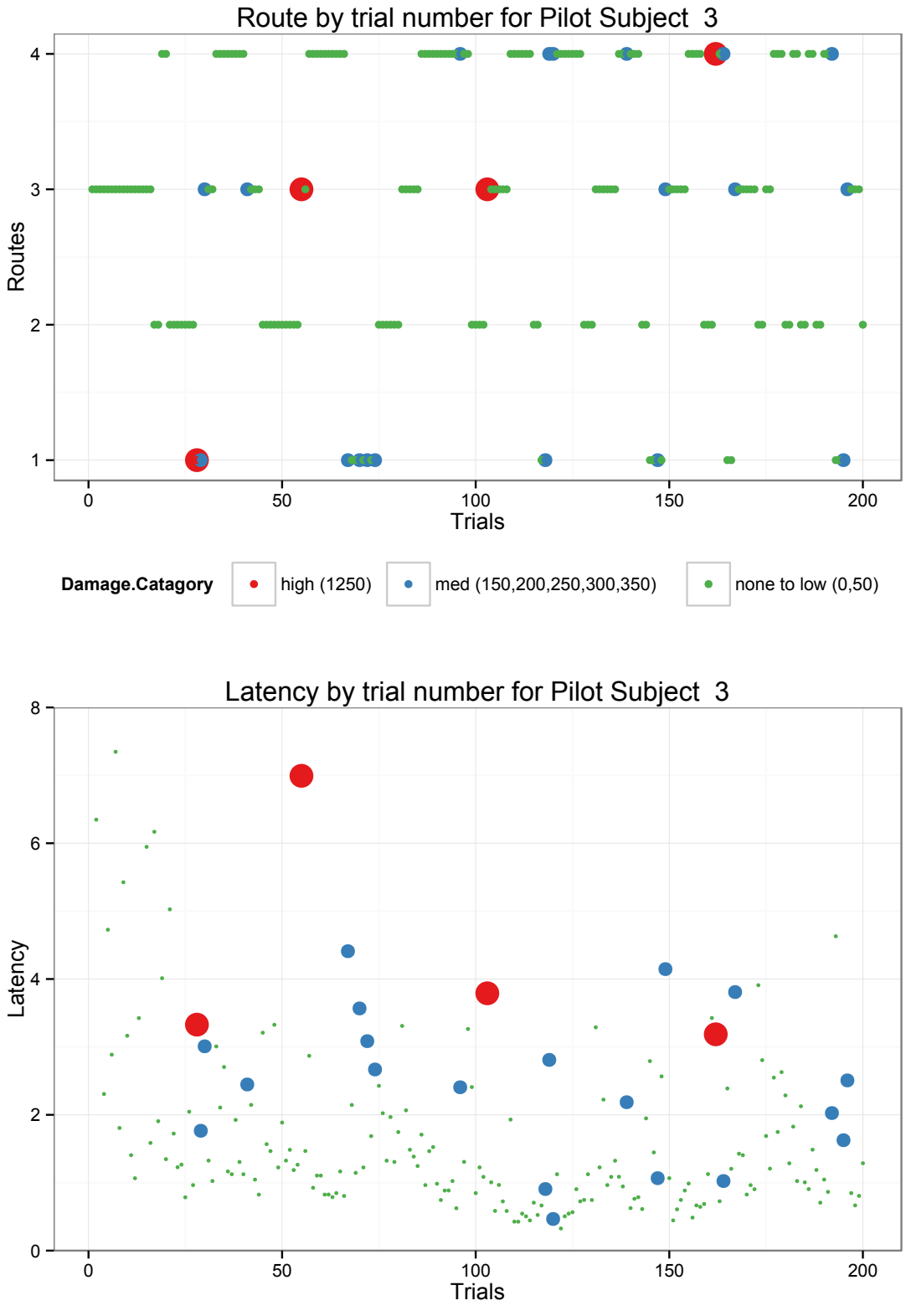


Figure 4.2: Route selection by trial for pilot subject 3 (top). Latency time by trial for pilot subject 3 (bottom). Damage.Category is the level of net damage received (Damage to Friendly Forces - Damage to Enemy Forces) on the previous trial. For example, the large red circles represent the next decision after receiving high friendly damage.

The graphs in Figure 4.1 tell us that pilot subject 1 was in exploration mode for about 10 trials; the longer latency times for the first 10 trials is consistent with exploration mode. Secondly, they tended to stick to the same route over several trials, even after receiving heavy damage. Finally, based on trial latencies they show a diminishing response to medium and heavy damage. In contrast, pilot subject 3, shown in Figure 4.2 ,tended to switch routes often particularly after receiving heavy damage. Their trial latencies reveal a clear response to medium and heavy damage throughout the task.

To determine if a greater number of trials would provide subjects with more opportunity to converge on an optimal decision pattern, Pilots 3 and 4 each completed 200 trials instead of the usual 100 trials. Below, we compare their results between their first and second 100 trials. A general pattern is found in which the pilot subjects made better decisions during the second half of the task than in the first half. The number of trials in which heavy friendly damage occurred either was maintained or decreased, the advantageous selection bias increased dramatically, and latency decreased. Additionally, pilot subjects showed a greater tendency to select the good routes (3 and 4). These preliminary results confirm our prediction that 100 trials may not be adequate for subjects to reach the exploitation phase. Therefore, in the actual study, 200 trials of the convoy task will be implemented.

	Pilot 3			Pilot 4		
Trials	1 - 100	101 - 200	1 - 200	1 - 100	101 - 200	1 - 200
Variables						
Final Damage	1550	500	2050	-50	2600	2550
# trials friendly damage	26	28	54	26	36	62
# trials heavy friendly damage	2	2	4	4	0	4
Route selection frequency						
Route 1	10	11	21	11	1	12
Route 2	27	21	48	42	6	47
Route 3	30	30	60	27	63	90
Route 4	33	38	71	20	31	51
advantageous selection bias	26	36	62	-6	88	82
mean latency time (sec)	4.110	1.367	2.738	3.668	.515	2.091
(SE)	(1.722)	(.095)	(.865)	(2.354)	(.033)	(1.180)
median latency (sec)	1.577	1.025	1.287	1.123	0.437	0.655
mode latency (sec)	1.025	1.226	1.125	0.655	0.374	0.374

Table 4.2: Descriptive statistics of pilot subjects 3 and 4 performance on convoy task comparing first 100 trials to second 100 trials. Note, Final Damage for trials 101-200 was calculated as the difference between damage on trial 200 and damage on trial 100.

Using the measure of performance regret discussed in section 2.4.1, we can compare the performance of the four pilot subjects. Figure 4.3 demonstrates the regret per trial for each pilot subject for the convoy task. The regret can be compared across subjects, we see a comparison of subject performance for the convoy task in Figure 4.4.

Figure 4.3: Regret for convoy task by pilot subject.

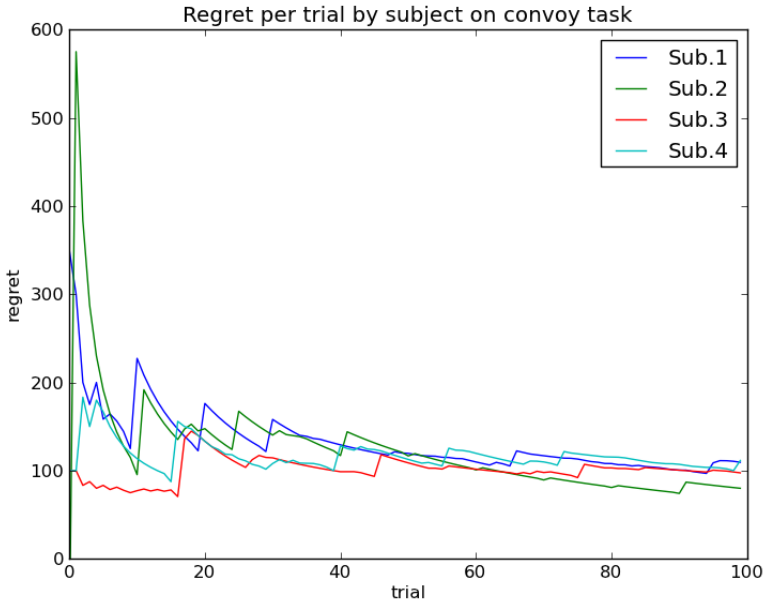


Figure 4.4: Consolidated regret, comparing subject performance for convoy task.

In sum, preliminary results from pilot data indicate that (1) the convoy task is a successful adaptation of the Iowa Gambling Task, (2) the high level of individual variability in the overall decision performance measures indicates that the need for the finer-grained exploration and exploitation measures as described in 2.4 above, and (3) extending the number

of trials from 100 (as in the IGT) to 200 will increase the likelihood of participants reaching the exploitation phase without causing undue participant burden.

4.1.2. Verification of Map Task

Results from the four pilot subjects demonstrate that our modification of the WCST provides ranges of decision scores consistent with healthy adult performance of the original WCST. Below, Table 4.3 contains descriptive statistics of many of the typical variables measured in the original WCST, along with latency times. Results also demonstrate that the map task elicits reasonable ranges of variability in the number of trials completed, percent of correct decisions, and the number of trials needed to complete the first sorting rule. An examination of the latency data reveals that, as expected, pilots subjects took more time to make their decision after making a wrong decision than when they had just made the correct decision.

Variable	Pilot 1	Pilot 2	Pilot 3	Pilot 4
# trials	86	79	98	115
% correct	88.37%	88.61%	81.63%	77.39%
Perseverative responses	6			
Perseverative errors	1			
% perseverative errors	1.163%			
Non-perseverative errors	0			
# trials to complete 1st category/rule	27	12	26	49
# categories achieved	6	6	6	6
Failure to maintain set	1	1	0	1
mean latency time (sec)	5.932	4.347	4.472	3.23
(SE)	(2.793)	(0.743)	(1.203)	(0.977)
	2.354	2.518	2.576	1.886
Mean latency previous trial **correct** (sec)	2.8121	3.103	2.669	2.014
(SE)	(0.238)	(0.246)	(0.11)	(0.099)
	2.239	2.422	2.352	1.811
Mean latency previous trial **wrong** (sec)	18.777	9.323	9.807	6.287
(SE)	(14.166)	(3.387)	(4.660)	(3.403)
	4.548	3.344	4.058	2.787

Table 4.3: Descriptive statistics of pilot subjects performance on map task. Definitions of variables can be found in Table 2.1.

4.2. Verification of EEG and Eye Tracking System

4.2.1. Verification of EEG

Approximately 25 iterations of EEG calibration across 10 people was conducted to ensure the EEG system would calibrate and provide good data during the decision-making tasks. Below are descriptive statistics of three EEG variables from pilot subject 4 during the convoy task: probability of distraction, probability of high engagement, and probability of cognitive workload. For each variable, higher values indicate higher levels of that particular cognitive state. These descriptive statistics indicate that pilot subject 4 was rarely distracted, and had moderate levels of engagement and cognitive workload during the task. Additionally, a reasonable range of cognitive workload occurred. CogState is ABMs general classification of type of brain activity and ranges from .1 to 1.0, as seen in Table 4.4. Scores of .3 indicate distraction, .6 low engagement, and .9 high engagement. As can be seen in Figure 4.5, 4.6 and 4.7, pilot subject 4 was predominantly in a state of high engagement, with few distractions.

Probability of Distraction		**Probability of High Engagement**		**Probability of FBDS (raw)Workload**	
Mean	0.04292	Mean	0.46328	Mean	0.63485
Standard Error	0.00593	Standard Error	0.01315	Standard Error	0.01154
Median	0.00003	Median	0.43721	Median	0.63989
Mode	0.00000	Mode	#N/A	Mode	#N/A
Standard Deviation	0.15625	Standard Deviation	0.34677	Standard Deviation	0.15526
Sample Variance	0.02441	Sample Variance	0.12025	Sample Variance	0.02411
Range	1.00000	Range	1.00000	Range	0.70874
Minimum	0.00000	Minimum	0.00000	Minimum	0.25140
Maximum	1.00000	Maximum	1.00000	Maximum	0.96014
Confidence Level(95.0%)	0.01164	Confidence Level(95.0%)	0.02583	Confidence Level(95.0%)	0.02277

Table 4.4: Subject 4 EEG data during convoy task.

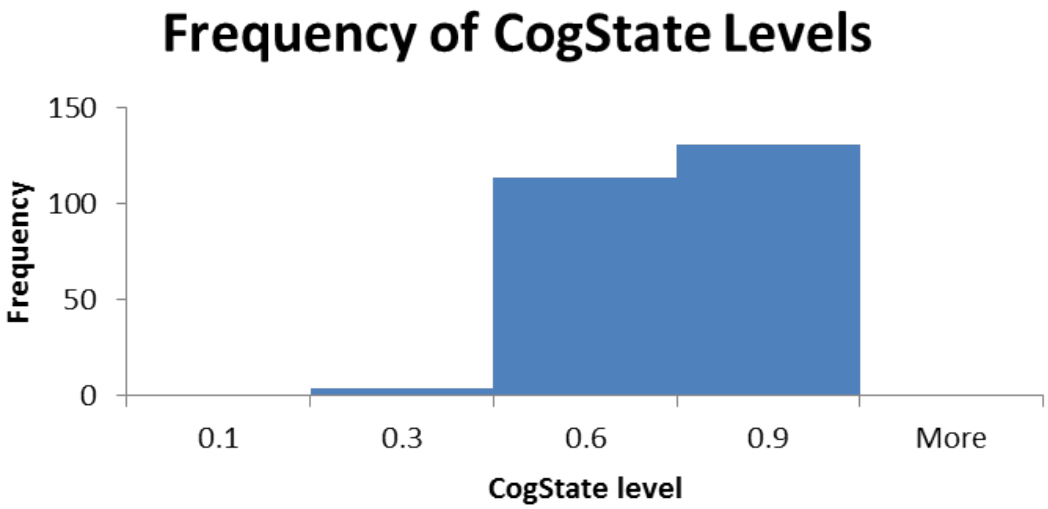


Figure 4.5: Subject 4 frequency of cognitive state levels.

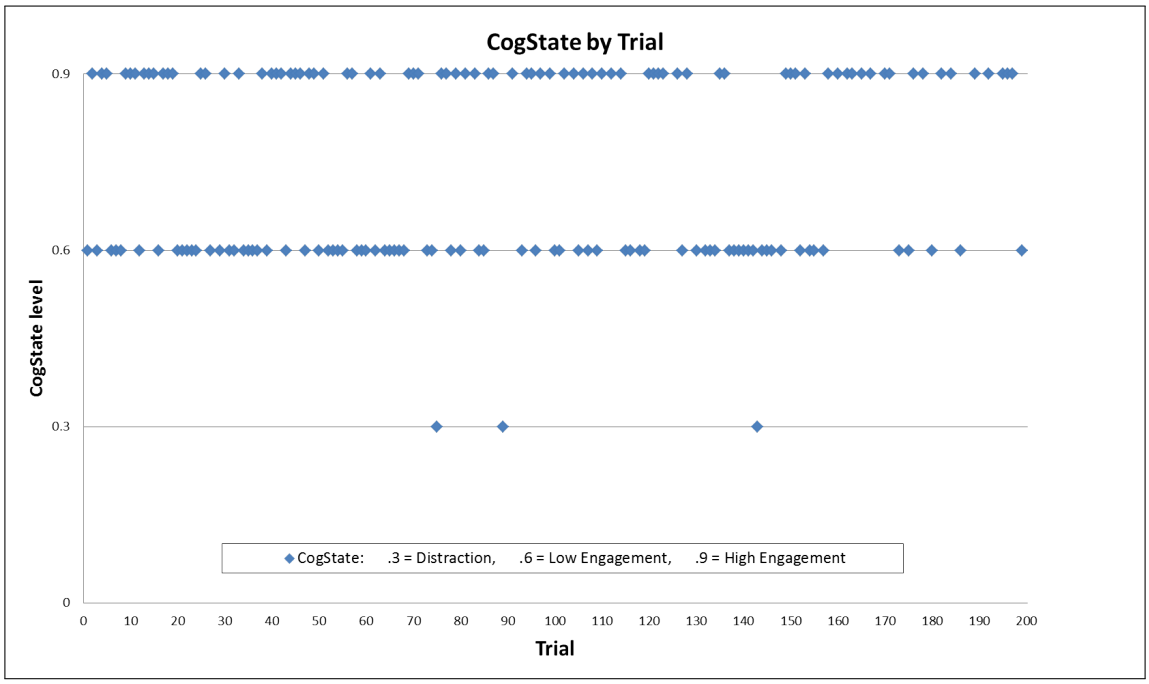


Figure 4.6: Subject 4 cognitive state levels by trial. No evident pattern as to when pilot subject 4 had high vs low engagement.

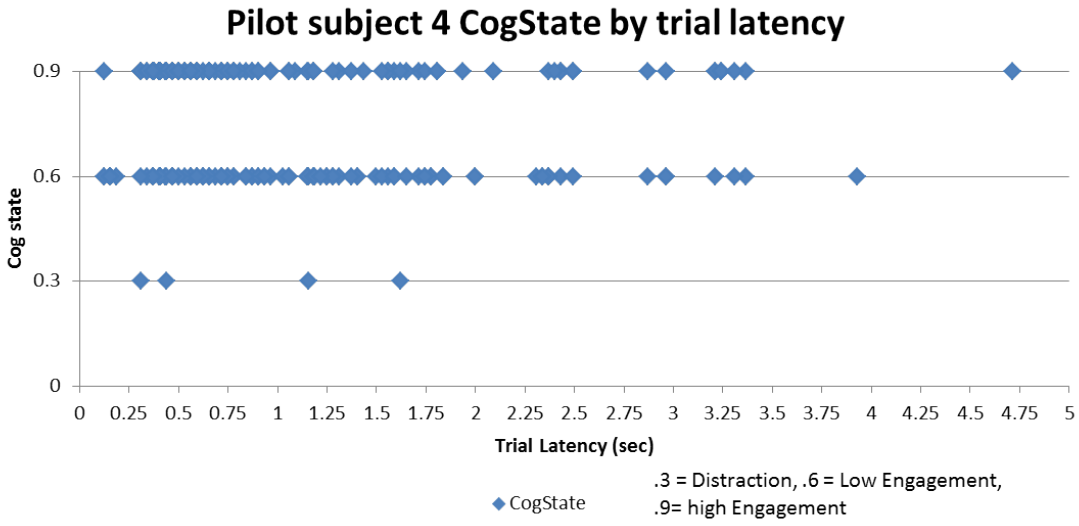


Figure 4.7: Subject 4 cognitive state levels by latency. No evident pattern that CogState is associated with trial latency.

4.2.2. Eyetracking

In the course of our pilot data collection, an unknown event occurred that caused the eye tracking software not to connect via the local network to the computer that collects EEG data. Various sources were consulted including NPS technical support and the developers of the software (Seeing Machines Inc. and EyeTracking Inc.), however the problem has persisted despite all efforts. To circumvent connectivity issues, we have purchased a computer with sufficient processing power to handle the EEG data collection software, FaceLAB eye tracking software, and the EyeWorks Record software to facilitate analysis. In the meantime, we show preliminary evidence that the eye tracking can accurately track visual scan as well as provide real time measurements of blink frequency, saccades and pupil diameter. Following are variables and their description used in eye tracking.

PERCLOS (percentage of eye-closure) PERCLOS is given as a percentage of measurement frames in a given time window where an eye was closed at least 75% of the way. A small number, close to 0, means eyes were mostly open. The time window is usually 10,000 frames (2 minutes, 46 seconds). This measurement does not include normal blinks.

SACCADE A saccade is a fast eye movement. Saccade value in the preliminary data analysis is the average of saccade observed for each data time. Smaller saccade average value means that eyes were fixated more than moving around: 1 = saccade, 0 = no saccade.

Pupil_diam The diameter of pupil dilation. The unit of pupil diameter is millimeters (mm).

Blink Freq The frequency of subjects blinking. The unit is Hz. For example, 0.2 means eyes blink 1 time every 5 seconds.

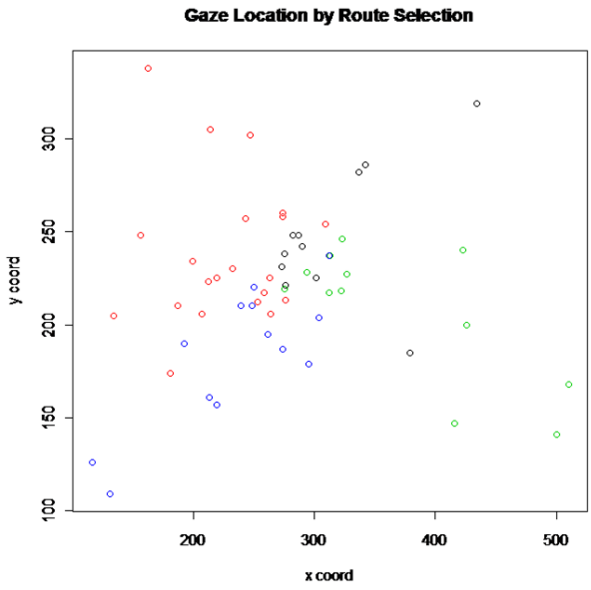


Figure 4.8: Gaze locations colored by route selection after the gaze.

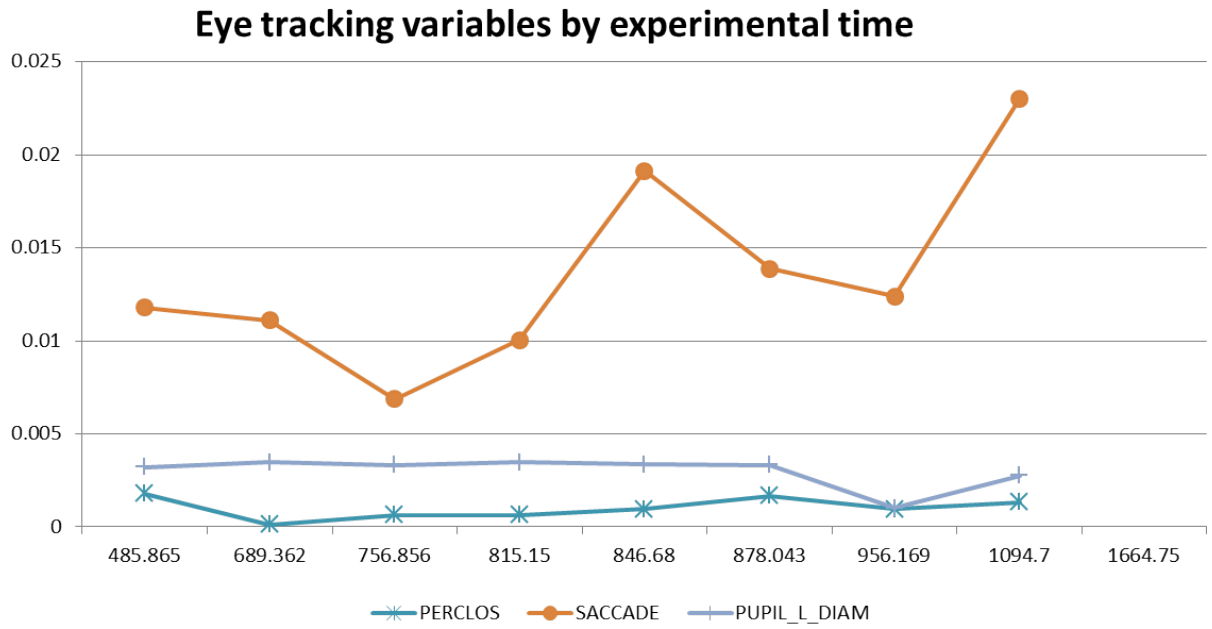


Figure 4.9: Eye tracking data from pilot sessions. The x-axis is experimental time, the y-axis is the average of the variable in each epoch/event.

5. STUDY 2 THESIS: A Comparison of Tactical Leader decision-making with Automated or Live Counterparts in a Virtual Environment (Virtual Battlefield Simulation 2)

The proposed thesis research will focus on determining if there is a difference between the decisions made by a leader of a Bradley vehicle section when their counterpart (wingman) is either a live being or an automated one. Decisions will be examined through the use of an Army Simulation, Virtual Battlefield Simulation 2 (VBS2). Subjects will be placed in a virtual environment where they will be required to make clearly defined tactical decisions. As part of the larger project, their brain activity and eye scan will be monitored via EEG and eye tracking. Their decisions will be evaluated and compared for accuracy, level of confidence, and time required. The primary research questions this research will answer are:

- When using a Bradley Section, do leaders decisions differ based on the type of wingman they are using - automated or live?
- Is there a difference in the amount of time for a leader to make a decision if using an automated system versus live units?
- Is there a difference in the leaders confidence levels when making a decision using an automated system versus US forces?

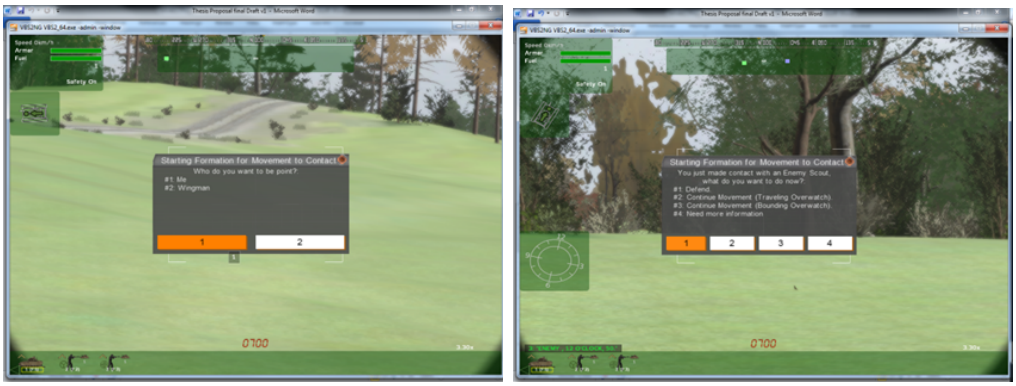


Figure 5.1: Examples of pop-up decision points encountered by subjects during study 2.

The benefits this study will provide are three fold. First, it will provide insight into the methods of tactical level decision-making that can be used by the Army Research Office for their study into how military personnel make decisions. Second, it provides a method for evaluating how an unmanned ground vehicle may be utilized in combat; specifically addressing the level of trust leaders now have in automated systems. Lastly, it provides validation for VBS2 as more than just a trainer, but also as a tool for conducting Simulation Based Acquisitions.

5.1. Methodology

The primary research questions will be addressed with a between subjects experiment. The experiment will comprise of approximately 30 subjects assigned to one of two groups, the artificial intelligence (AI) group and the live group. In the AI group, participants will be instructed that they will use an AI. In the live group, participants will be instructed that they have a live crew counterpart. Subjects will be active duty military personnel from the US Army and Marine Corps. The groups will be blocked by their experience level, as determined by a pre-test of basic tactical knowledge required for the scenario and use of Unmanned Systems. Both groups will conduct the same scenarios in VBS2 except with slightly different interfaces which add to the perception of a responsible AI or a live crew.

Throughout the 10-20 minute scenario, several pop-up decision points will occur. The subject will not be able to proceed without making a decision. There are two basic types of decisions, Tactical Decision and Movement Decision. Each Movement Decision comprises two possible options, to either let the UGV lead or the subject lead. Tactical Decisions will have up to four options, one of which will be a request for more information. This option allows the subject to look over their map or through their screen resources prior to making a decision. The other three decisions will be used to provide decisions about tactical actions within the scenario. The decisions chosen by the subject will be recorded each time they make a decision, as will their level of trust in the AI or live counterpart. As part of the larger project, the use of EEG and eye eye tracking equipment will be used to facilitate understanding about the subjects cognitive load, what information they looked at, and their degree of distraction. In addition, during the study, the evaluator will ask the subject how confident they were in each of their decisions immediately following the decision.

Both groups will be given the same decision points and possible paths so that a comparison can be made between the mean of each groups path choice. The decisions made by the subjects will be evaluated based on their path score. Appropriate statistical methods, such as a 2-sample test, will be used to test the hypothesis.

To facilitate recording of path choice and determining a mean path, 96 possible path choices are outlined in Appendix D.12, providing an overall path score for each trial. All movement decisions are two level, whereas the first tactical decision is three level and the second tactical decision is two level. It is important to note, that at each tactical decision, there is one additional option, a request for more information. The subject will be able to choose more information a maximum of four times. The use of the more information option will not get factored into the decision point value. The number of times more information is used, however, does account for the total path score. The path score, then, is calculated by adding the values for the decisions to the information score. This provides a distribution of the path score that only allows one complete success and one complete failure, with the majority of results dispersed to the median of the total paths. A post scenario survey will ask subjects to explain their decisions, allowing for greater insight as to why they chose their actions.

5.2. Neurophysiological Model of Tactical Decisions

Because of key aspects of the decision task, the combination of real-time neurophysiological and behavioral decision data will extend upon our understanding of optimal wargaming decision-making: the task is dynamic, captures real-world tactical decisions and participants are provided with a mix of relevant and irrelevant visual information. Additionally, results will provide insight into how tactical leaders handle new technology (such as an automated wingman). With these characteristics, we will be able to test the hypothetical model of dynamic decision-making described in Introduction. For ease of reference, the model is shown again in Figure 5.2.

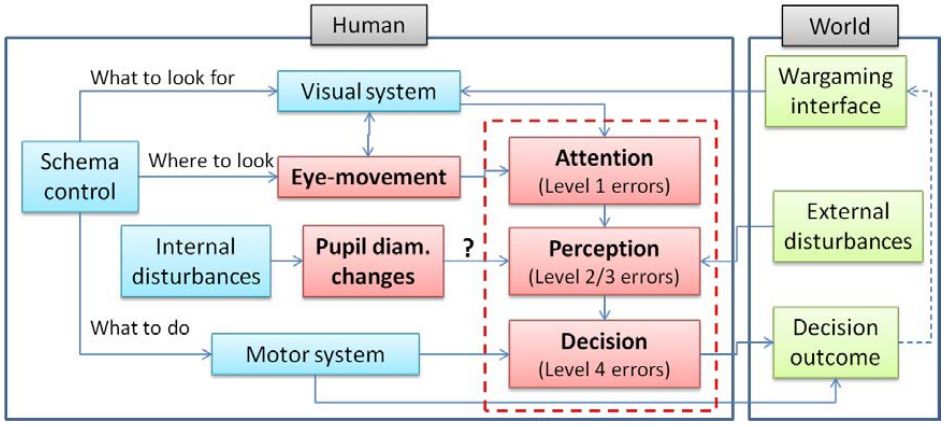


Figure 5.2: Proposed hypothetical structure of decision-making considering neural system, gaze control, and world.

6. CONCLUSION

6.1. FY13 Progress

The following items generally list the measures of progress towards research project completion.

- *IRB approval from NPS and ARO received.* Appendix C.1 for documentation of IRB approval.
- *Development of wargames to observe military decision-making.* The convoy task, developed from the IGT, simulates military resource allocation (convoys) with variable payoffs for investment (damage to friendly/enemy forces). The map task, developed from the WCST, simulates rule development (sorting maps with graphic control icons) and re-evaluation in the face of changing schedules of reinforcement.
- *Successful implementation of EEG.* Added LT Lee Sciarini (Dr.) to the team as an EEG consultant. Acquired synchronization software to synchronize EEG, eye tracking and behavioral data. Successful implementation of EEG in terms of calibration and collection of 'good' data.
- *Pilot data collection.* 4 pilot subjects completed finalized versions of each task while EEG data collected. Pilot testing indicated an adequate range of decision performance is captured for both the convoy and map task.
- *Initiation of study 1.* Recruiting subjects for study 1 (convoy and map tasks) data collection.
- *Recruitment of thesis student.* MAJ Scott Patton traveled for VBS2 training to develop tactical wingman task for study 2.
- *Proposed statistical method* Using sequential detection methods discussed in Chapter 2.5, we propose to identify when a subject switches from *exploration* to *exploitation*.
- *Presentation to Rear Admiral Doll on Aug 6th.* During a site visit to the Naval Postgraduate School Naval Postgraduate (NPS) institutional review board (IRB), RADM Doll received a brief on this project as a representation of the research conducted at NPS. He was impressed with the work and discussion focused on human subject experimentation and IRB considerations, complexities of military decision-making, and data analysis challenges and opportunities.
- *Project meetings.* In the course of meeting objectives, the team completed 30 weekly and 8 monthly meetings.

6.2. Initial Findings

The Pilot Data Collection demonstrates that:

- Range of behavior performance is acceptable to move forward.
- Preliminary evidence that the balance between exploration and exploitation can be captured.
- Experimentation procedures are on track to allow a synchronization of eye tracking, EEG and behavior.
- The data collected has shown promise in revealing patterns for EEG and eye tracking.
- The high level of between subject variability in decision performance speaks to the need for the proposed decision models.

6.3. Future Work

For the second year, we are about to start conducting study 1. This will include data collecting, cleaning and analyzing. We also will begin to write up results for submission to peer reviewed journals and conferences. We anticipate designing a follow-on study based on Study 1 results. We will continue to recruit additional thesis students. The thesis study will be conducted and completed in FY14.

Potential scholarly paper topics from studies one and two include:

- Demonstration of successful modification of IGT and WCST into military relevant decision-making tasks.
- Correlation between neurophysiological measures and decision performance.
- The role of working memory and visual processing speed on military decision-making.
- Modeling human decision-making on modified IGT and WCST (method of maintaining estimate, level of exploration, level of discounting).
- Comparing performance of algorithms on modified IGT and WCST.
- Assessing decision-making performance with EEG to guide training interventions.
- Comparing decisions and underlying cognitive strategies differ when tactical leaders work with a live wingman versus an automated wingman.

For year three, we expect to complete papers from study 1 and 2 and to conduct and report the results from the follow-on study designed in year two.

APPENDIX A. Project Methodology

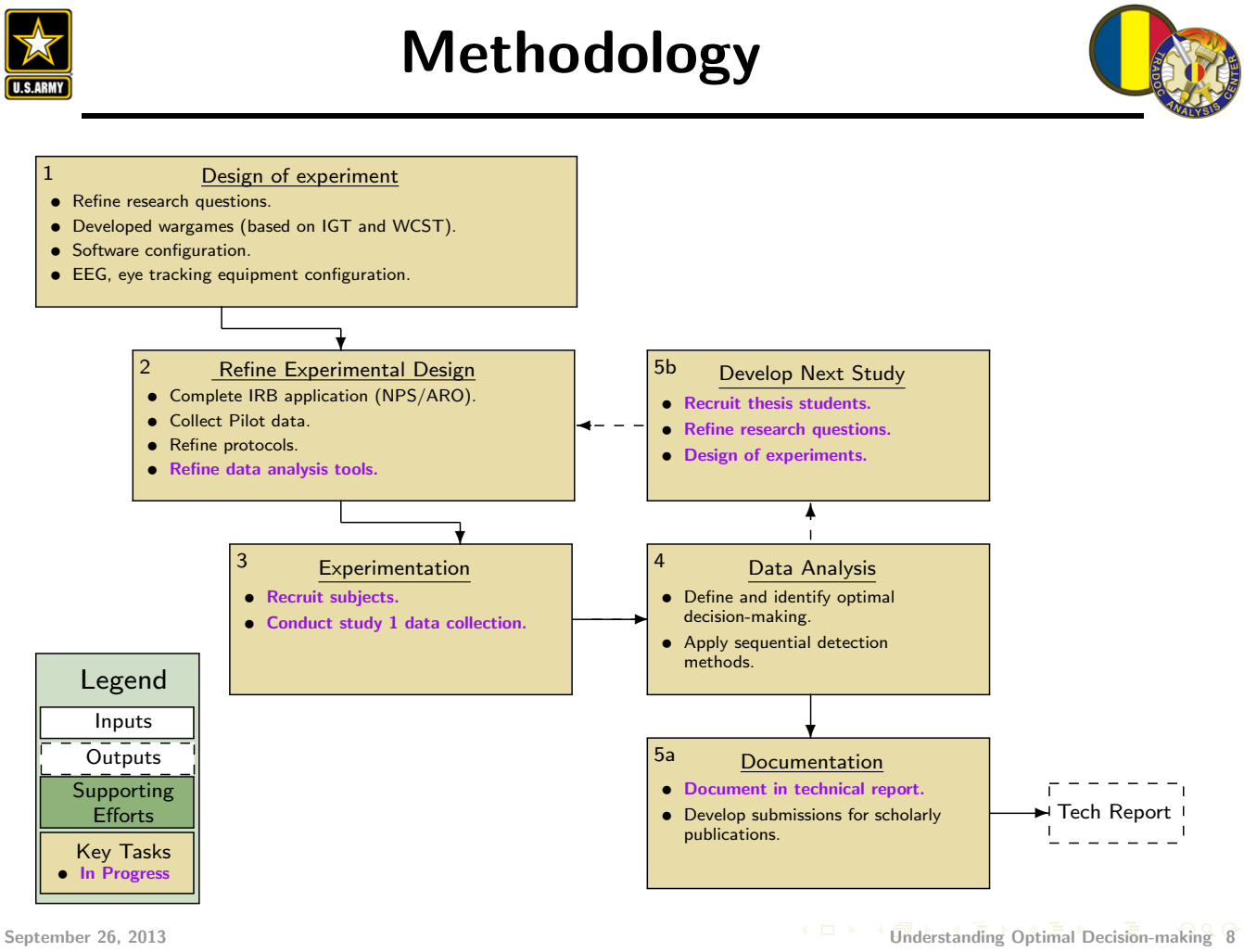


Figure A.1: Methodology flowchart for Project 638, Understanding Optimal Decision-making in Wargaming from Close of FY13 IPR brief, 26 September 2013.

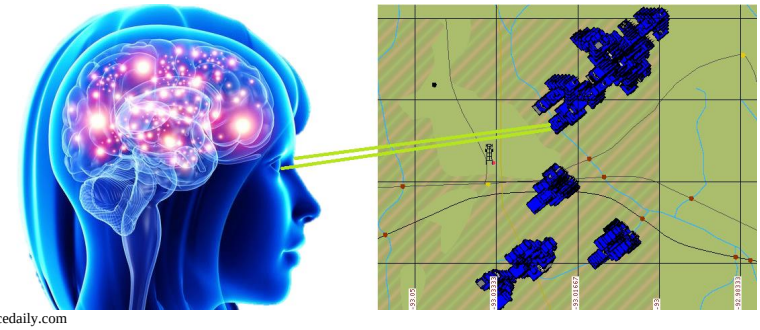
• Accelerating research results transition to applications in all stages of the research and development cycle.
• Strengthening academic, industrial, and nonprofit laboratories research infrastructures which serve the Army.
• **Focus on those research topics that support technologies vital to the Army's future force, combating terrorism and new emerging threats.**
• **Directing efforts in research areas relating to new opportunities for Army applications and which underscore the role of affordability and dual-use, especially as they provide new force operating capabilities and emerging threats.**
• **Leveraging the science and technology of other defense and Government laboratories, academia and industry, and appropriate organizations of our allies.**
• **Fostering scientist and engineer training in the disciplines critical to Army needs.**
• Actively seeking creative approaches to enhance education and research programs at historically black colleges and universities and at minority institutions.

Table A.1: Army Research Office Functions (ARO, 2012).

APPENDIX B. Recruitment Advertisement

Military Decision Making Study
Volunteers needed.

Come test your decision making skills and get involved in cutting edge research. Be a single visit volunteer for our military decision making study taking place in TRAC-Monterey (in Watkins) sponsored by the Army Research Office and NPS.



Sciencedaily.com

You will be asked to complete some military decision making tasks while your eye gaze and brain activity is monitored via eyetracking and EEG technology. The purpose of the study is (1) to test the validity of newly created military decision making tasks; (2) to attempt to characterize military decision making through the use of behavioral and neurophysiological measures.

WHO is eligible: Active duty military students, faculty and staff.
WHERE: TRAC-Monterey Battle Simulation Lab, Watkins Building, Room 191
HOW LONG: Approximately 2 hrs.
WHEN: You can choose your own experiment time.
CONTACT: Mr. Jesse Huston at jesse.huston@gmail.com to schedule.

Risks associated with this study are minimal. Participation is completely voluntary. The principal investigator of the study is Dr. Quinn Kennedy (mqkenned@nps.edu). Please contact NPS IRB Chair Dr. Larry Shattuck (lgshattu@nps.edu) with any questions regarding your rights as a participant.

Figure B.1: Advertisement for recruitment of subjects.

APPENDIX C. IRB Approval of Study 1 Protocol Memorandum

Naval Postgraduate School
Human Research Protection Program

AUG 19 2013

From: Interim President, Naval Postgraduate School
Via: Chairman, Institutional Review Board
To: Dr. Quinn Kennedy, Operation Research Department
LTC Jon Alt, TRAC-Monterey
MAJ Pete Nesbitt, TRAC-Monterey
Lee Whitaker, Operation Research Department

SUBJ: UNDERSTANDING THE DEVELOPMENT OF OPTIMAL DECISION MAKING

Encl: (1) Approved IRB Protocol

1. The NPS IRB is pleased to inform you that the NPS Interim President has approved your project (NPS IRB# NPS.2013.0066-AM01-EP7-A). The approved IRB Protocol is found in enclosure (1). Completion of the CITI Research Ethics Training has been confirmed.

2. This approval expires on 21 July 2014. If additional time is required to complete the research, a continuing review report must be approved by the IRB and NPS President prior to the expiration of approval. At expiration all research (subject recruitment, data collection, analysis of data containing PII) must cease.

3. You are required to obtain documented consent according to the approved procedure provided in the approved protocol.

4. You are required to report to the IRB any unanticipated problems or serious adverse events to the NPS IRB within 24 hours of the occurrence.

5. Any proposed changes in IRB approved research must be reviewed and approved by the NPS IRB and NPS President prior to implementation except where necessary to eliminate apparent immediate hazards to research participants and subjects.

6. As the Principal Investigator (PI) it is your responsibility to ensure that the research and actions of all project personnel involved in this study will conform with the IRB approved protocol and IRB requirements/policies.

Figure C.1: Naval Postgraduate School IRB approval of study 1 protocol memorandum.

SUBJ: UNDERSTANDING THE DEVELOPMENT OF OPTIMAL DECISION MAKING

7. At completion of the research, no later than expiration of approval, the PI will close the protocol by submitting an End of Experiment Report, all signed informed consent forms and research data on a CD (including audio recordings and research notes) to the IRB for storage. The IRB will secure these documents for 10 years and then forward to the nearest FRC.

Lawrence G. Shattuck, PhD
Chair
Institutional Review Board

RDML Jan E. Tighe, USN
Interim President
Naval Postgraduate School

Figure C.1 (continued): Naval Postgraduate School IRB approval of study 1 protocol memorandum.

APPENDIX D. Testing Procedure Tools

D.1. Consent to Participate Form

Naval Postgraduate School
Consent to Participate in Research

Introduction. You are invited to participate in a research study entitled "Understanding the development of optimal military decision making." The purpose of the research is (1) to test the validity of newly created military decision making tasks; and (2) to attempt to characterize the development of optimal military decision making among military personnel through the use of behavioral and neurophysiological measures.

Procedures. We are asking approximately 30 military personnel to complete two military decision making tasks while your eye gaze and brain electrical activity are monitored via eyetracking and EEG technology. In one task, you will see four roads displayed on a computer monitor. Over a series of trials, you must select the road your convoy should use. In the other task, you will see five digital representations of command and control maps displayed on a computer monitor. Over several trials, you match one of the four maps displayed at the bottom of the screen with the map displayed at the top of the screen. You also will be asked to complete a demographic survey, a visual acuity test, some brief cognitive tests, and a post-task survey. The expected duration of your participation is approximately 2 hours. These procedures are new and are only related to the research and serve no purpose other than this research endeavor.

Location. The study will take place at TRAC-Monterey.

Cost. There is no cost to participate in this research study.

Voluntary Nature of the Study. Your participation in this study is strictly voluntary. If you choose to participate you can change your mind at any time and withdraw from the study. You will not be penalized in any way or lose any benefits to which you would otherwise be entitled if you choose not to participate in this study or to withdraw. The alternative to participating in the research is to not participate in the research.

Potential Risks and Discomforts. We anticipate no to very minimal discomfort. Participants will sit in front of a LCD computer monitor and two small cameras while they complete computerized tasks. The eyetracking camera systems have infrared lights for gaze detection purposes. The infrared lights are no more harmful than normal room lighting. The EEG system is wireless and runs on a battery, so risk of physical discomfort is practically nonexistent. Electrolytic gel consists mostly of saline solution and easily washes off of skin and out of hair. There is no to very minimal likelihood that you will experience any emotional discomfort from any of the decision making or cognitive tasks.

Anticipated Benefits. Anticipated benefits from this study are greater insight into the neurophysiological underpinnings of the development of optimal decision making. Additionally, the study tests the feasibility of transforming standard psychological tests into tasks that tap common military decisions. Therefore, the decision making tasks have the potential to be used in a wide range of military settings/applications attempting to better understand and/or facilitate optimal wargaming decision making. You may gain insight into how you develop decision making strategies.

Compensation for Participation. No tangible compensation will be given.

Confidentiality & Privacy Act. Any information that is obtained during this study will be kept confidential to the full extent permitted by law. All efforts, within reason, will be made to keep your personal information in your research record confidential but total confidentiality cannot be guaranteed. In accordance with NPS data storage instruction, data only will be kept on approved NPS systems. Hard

Version #
Date:

Figure D.1: Naval Postgraduate School IRB approved consent to participate in research.

Consent to Participate Form (continued)

copies of informed consent forms and demographic data will be kept separately in a locked cabinet in a locked laboratory. Hard copies of data will be associated only with the participant identification number.

If you consent to be identified by name in this study, any reference to or quote by you will be published in the final research finding only after your review and approval. If you do not agree, then you will be identified broadly by discipline and/or rank, (for example, "fire chief").

☐ I consent to be identified by name in this research study.

☐ I do not consent to be identified by name in this research study.

Points of Contact. If you have any questions or comments about the research, or you experience an injury or have questions about any discomforts that you experience while taking part in this study please contact the Principal Investigator, Dr. Quinn Kennedy, 656-2618, mqkenned@nps.edu. Questions about your rights as a research subject or any other concerns may be addressed to the Navy Postgraduate School IRB Chair, Dr. Larry Shattuck, 831-656-2473, lgshattu@nps.edu.

Statement of Consent. I have read the information provided above. I have been given the opportunity to ask questions and all the questions have been answered to my satisfaction. I have been provided a copy of this form for my records and I agree to participate in this study. I understand that by agreeing to participate in this research and signing this form, I do not waive any of my legal rights.

_____ _____
Participant's Signature Date

_____ _____
Researcher's Signature Date

Version #
Date:

Figure D.1(continued): Naval Postgraduate School IRB approved consent to participate in research survey.

D.2. Procedure Notes

Before the Subject Arrives

D.2.0.1. After recruitment

- Obtain nasion to inion *(front to back)* and helix-helix *(ear to ear over top)* head measurements and determine whether the subject needs a small or medium electrode strip. Use the measurements to refer to the chart on the laminated sheet. *The midpoint between nasion and inion is the Cz point on the electrode.*
- Inform the subject that they will need to be able to perform this test WITHOUT GLASSES. *Contacts are O.K.! (20/30 vision is tolerable)*

D.2.0.2. Before the subject gets there

- Turn on the facelab laptop and update its antivirus software (this will allow it to transmit data over the network). Open facelab 5.0 and use the single configuration option that is available. Run Symantec LiveUpdate on the desktop computer as well. *Due to recent security measures taken by NPS, network activity is BLOCKED on systems that do not have up-to-date anti virus software. This will prevent the eye tracking system from being able to communicate with the desktop.*
- Determine which electrode strip to use (see above).
- Put electrode pads (with blue tags) onto EEG recording sites. The pads do no recording, but rather house the conductive gel that allows the electrodes to take measurements.
- When applying gel to electrodes, fill the center of the pad with gel directly and add a modest amount to the area around the hole at the top to help with saturation. Start filling from the bottom and work your way up. It is possible for bubbles to form and for the pad to absorb some gel.
- Prepare two adhesive electrodes for impedance reference. These will be applied to the mastoid bones behind the ears.
  - Trim the white adhesive pads of the electrodes so that they will be able to lie flat against the subject's skin.
  - The metal portion is not adhesive, but is the area that records the data. Apply gel to the metal portion. Enough so that it will make contact with the skin, but not enough so that pressing on the electrode will cause the gel to leak into the adhesive.

Verify Calibration

- On the control window, select the verify calibration button. Hold the calibration key with one handle in each hand so that the air bubble is facing upward. Allow three snapshots to be taken. Do not move the key while the you head a shutter closing sound effect. Do not move the key like you would a steering wheel.
- If you are told that they eye tracker does not need to be calibrated, proceed to making a head model. If you are told that the eye tracker does need to be calibrated, proceed to the next step.
- For a more detailed version, follow the onscreen instructions. Otherwise:
  - Choose your environment (laboratory).
  - Choose precision mode.
  - Make sure both cameras can see the subject.
  - Skip next screen.
  - Using the large image, adjust the individual cameras aperture size and focus to eliminate graininess and improve clarity. Click Next to do the same for the next camera.
  - Use the calibration key to take snapshots so that they program understands the orientation of the cameras.
  - Allow the key to hang from your thumbs.
  - Try to make the key fill up as much of each cameras visual field as possible.
  - As long as the key is visible, the cameras will take snapshots at regular intervals, try not to move the key while the shutter is closing.
  - Take pictures of the key straight on, slightly to the right, left, upward, and downward.
  - DO NOT rotate the key like a steering wheel.
- Take a snapshot of the calibration key, making sure that the level has the bubble in the middle.
- Select the option to calibrate using a key. There should be a USB dongle attached to the laptop for this.

Otherwise, Calibrate cameras if you:

- Have put the camera in a new place (vertical rotation of both cameras at once does NOT require recalibration).
- Have moved either of the cameras relative to one another.

- Changed the facing of the cameras.
- Are unsure about whether you should recalibrate the cameras.
- Are faced with a subject whose eyes just wont track well.

Forms

- Write the subject ID number and date on their demographic survey, post-task survey, trails A & B.
- Do NOT write the subjects name on any of the pieces of paper with their ID number.
- Do NOT write the subject ID number on the informed consent form.
- Get a stopwatch for trails A and B. This can be a stopwatch function on a smartphone or online (e.g. www.online-stopwatch.com).

Meet and Greet

Script (say this exactly): You are here for a military decision-making study. First we will go through some paper and pencil tests. We will then move on to work with the computer. While you are doing the computer tasks, we will be monitoring your eye gaze patterns with eye tracking technology and your brain activity with EEG. Please note that this study is completely voluntary and that you may choose to opt out of it at any time. Should you choose to opt out, there will be no repercussions.

Forms and questionnaires

- Give the subject the informed consent form. No other forms may be given until this form is signed and dated. Give the subject an additional unsigned copy of the consent form that they may keep for their records.
- Once the consent form has been signed, give the subject the demographic survey. Check to make sure that they are active military, else they must be excluded from the study.

Snellen Test

- In the battle lab, behind the door, there is a Snellen eye chart. Close the door and have the subject stand behind the masking tape on the floor near the cubicles with their toes touching the tape.

- Have the subject obscure their right eye. Tell the subject NOT to press on the eye they are obscuring. DO NOT GIVE HINTS. Ask them to read the 20/30 line. Close answers are not acceptable (e.g. P for F). The subject may have at most one wrong answer for a line to be considered to have that line correct. If the subject gets a line wrong, move upward on the chart until they are able to get a line completely correct (or with one miss). If the subject gets a line correct, move down the chart until they give two or more incorrect letters. Record the answer for the left eye with the highest acuity on the demographic form. To ensure that subjects to not memorize answers, once you have recorded their acuity with an eye, have them read two or three additional random lines.=
- Repeat the process for the opposite eye, then repeat the process with both eyes open.

Trails A

- Script: I will give you a sheet of paper on which you will see numbers from 1 to 25. Your job will be to draw a line from 1 to 2, 2 to 3, and so on until you reach the end of the numbers. Work as quickly and accurately as you can. I will tell you when to start. I will now give you a demonstration. DO NOT TELL THEM TO START. Do the trails A short demo so that they can see.
- Get the stopwatch, prepare to time the subject (clear the stopwatch to zero). If the subject makes a mistake, IMMEDIATELY correct them.
- Answer any questions the subject has, then administer trails A while timing them.
- Record their time on the sheet.

Trails B

- Script: I will give you a sheet of paper on which you will see letters and numbers. Your job will be to draw a line from 1 to A, A to 2, 2 to B, and so on until you reach L. Work as quickly and accurately as you can. I will tell you when to start. I will now give you a demonstration. DO NOT TELL THEM TO START. Do the trails B short demo so that they can see.
- Get the stopwatch, prepare to time the subject (clear the stopwatch to zero).
- Answer any questions the subject has, then administer trails B while timing them.
- Record their time on the sheet.

EEG procedure notes

Applying the EEG:

- Wipe the subject's hair down lightly with an alcohol swab or use a comb dipped in rubbing alcohol before applying EEG.
- Attach the neoprene strap to the B-Alert headset and find a comfortable size for the subject. Make sure that the strap is even on both sides so that the triangular portion of the strap is directly opposite the center of the headset.
- CHECK: the reference electrode leads should be pointed downward.
- CHECK: the strap should be tight enough that it will not easily move and will rest ABOVE the ear lobes. The strap should also be loose enough that the subject will not find it distracting.
- Locate the mastoid bone directly behind the meat of the ear. Place one of the prepared adhesive metal electrodes directly on the hard mastoid bone. Trim the white adhesive portion of the electrode with scissors to help it fit on the mastoid. Take care NOT to place the metal portion over any muscle (which is squishy). Once these electrodes have been placed, clip the wires from the headset onto them.
- Apply the cap to the scalp by placing the triangular portion of the neoprene strap through the hole on the electrode strip that is opposite the strips plug. To help with the connection, make sure that the strap is on the outside of the electrode cap and that the triangle moves in toward the scalp.
- Tips for applying gel to sensors and decreasing resistance:
  - Note where gel makes contact with hair or scalp and use that as a reference point when applying more gel.
  - Use the gel syringe, tweezers, or the handle end of the comb (sterilized with alcohol) to move hair away from points where the electrode pads meet the scalp.
  - While the pad is making contact with the head, you may squeeze the syringe tip under the pad. When doing this, lightly press on the plunger and rotate the tip so that it turns directly from the pad to the scalp. This will help make contact.
  - Make sure the electrode cap is fastened tightly to the strap. This will help the cap stay on the subjects head and will greatly reduce the impedance in the signal.
- Apply more gel to areas where the gel from electrodes leaves residue on the head. The goal is to provide a direct gel channel from the scalp to the electrode. This is the crux of the recording process.

Starting the software

- BEFORE OPENING THE VISUAL SOFTWARE, make sure that you turn the headset on and that it is showing a solid green LED and no yellow LED. This should also mean that the ESU or dongle is showing a solid light.
- Important: plug the electrode strip into the headset before testing impedances. Test impedances first. In order to get a reading on the reference electrodes, the cap MUST be plugged into the headset. It is imperative that the reference electrodes have resistances in the green range of values before worrying about impedances on other sites. Reference electrodes affect resistance levels on all other sites.
- If an electrode gives impedance values above 40 kOhms, lift the electrode and do any or all of: clean the site, apply more gel, remove excess hair, ensure that contact between skin, gel, and electrode is indeed made.

EEG Baselines

D.2.0.3. 3CVT, EO, EC

Run the baseline tasks on the EEG software. This should take approximately 15 minutes total. The tasks are the three-choice vigilance task (3CVT), eyes open (EO), and eyes closed (EC). F3 skips to the end of a particular baseline task, but may only be done when the instructions are shown. F8 interrupts a task, allowing a user to continue, restart a practice, restart testing, or skip to a new task. F11 exits from the, Please wait for the technician..., Technician assistance requested, and Thank you for completing... windows.

D.2.0.4. Digit Span

- Non-baseline task. Do the practice version of the test.
- "The researcher is going to say some numbers. Listen carefully. He can only say them one time. When he is through, please say them back to him in the same order. Just say what he says."

D.2.0.5. Reverse Digit Span

- SEE ABOVE for Digit Span.
- The researcher is going to say some more numbers, but this time when he stops, please say the numbers backward. If he says 7-1, what would you say? Let's try another: 3-4.

Calibrate Eye Tracker

Making a Head Model

- File, Create New Model (Ctrl + N)
- Have subject look at eye level in the direction of the cameras.
- Take a snapshot. The subject may move his/her head freely at this point.
- Adjust feature markers to the corners of the eyes and the corners of the mouth.
- Check that the feature tracking responds to the proper areas.
- Decide between iris and pupil tracking.
- Set the desired saccade threshold.
- (Optional) Calibrate the gaze by having the subject look into the right camera, then the left when prompted.

D.3. Running the experiment:

- Make sure the subject is not fidgeting or resting their face on their hand.
- The subject MAY NOT chew gum.
- Sniffling or fidgeting due to illness or allergies may generate artifacts that complicate the interpretation of the data.

Post-Task Survey

- Ask the subject to complete the post-task survey. Answer whatever questions they have.
- Mention that the subject should check their head in the mirror for gel. It is water soluble.

Post Task Cleanup:

- Let the subject remove the reference electrodes from their mastoid bones.

- Turn off the EEG headset and undo the neoprene strap. Then remove the strip and strap from the subjects head carefully (hair may get caught in the velcro).
- Be sure to clean the EEG after every use. Use the tweezers to remove electrodes and ensure that all sites have all of their adhesive removed. Dispose of used mastoid electrodes and sensor pads in the garbage. To remove excess adhesive and gel, use a cotton ball dipped in distilled water and dry off the site afterward. Allow the electrode strip to dry.

D.4. Demographic Survey Sheet

ARO Study 1
Demographic Survey

Subject # Date

1. Age:_____
2. Gender: Male_____ Female_____
3. What is your preferred hand for writing? Right_____ Left_____
4. Do you serve or have you served in any armed forces? Yes No
5. If yes, which branch? _____ Rank:_____ Years:_____
6. How many total months have you been deployed?
7. When was your most recent deployment?
8. Where was your most recent deployment?
9. During your most recent deployment, what were your main responsibilities?

****

To be completed by the experimenter:

Visual acuity:
Left eye _____
Right eye _____
Overall _____

Figure D.2: Demographic survey used before the Pilot Test.

D.5. Snellen Eye Acuity Test Chart

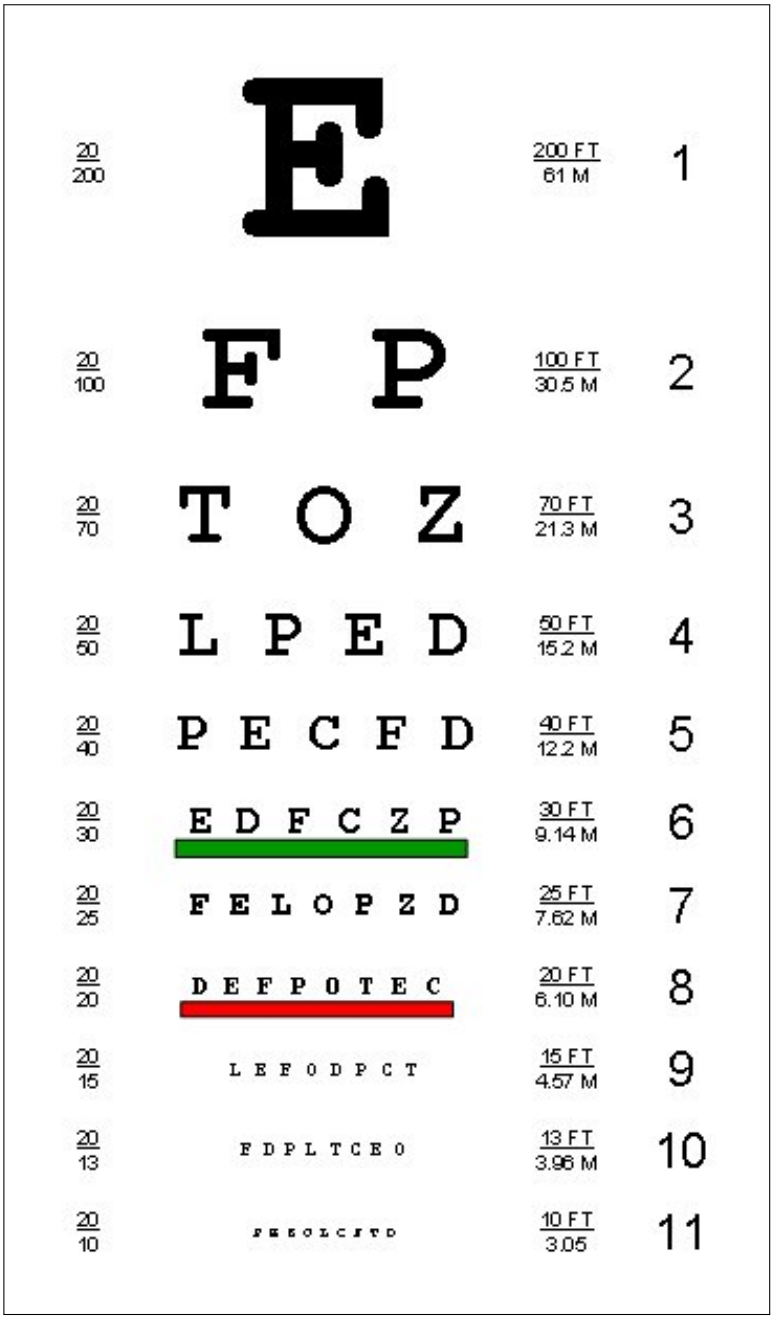
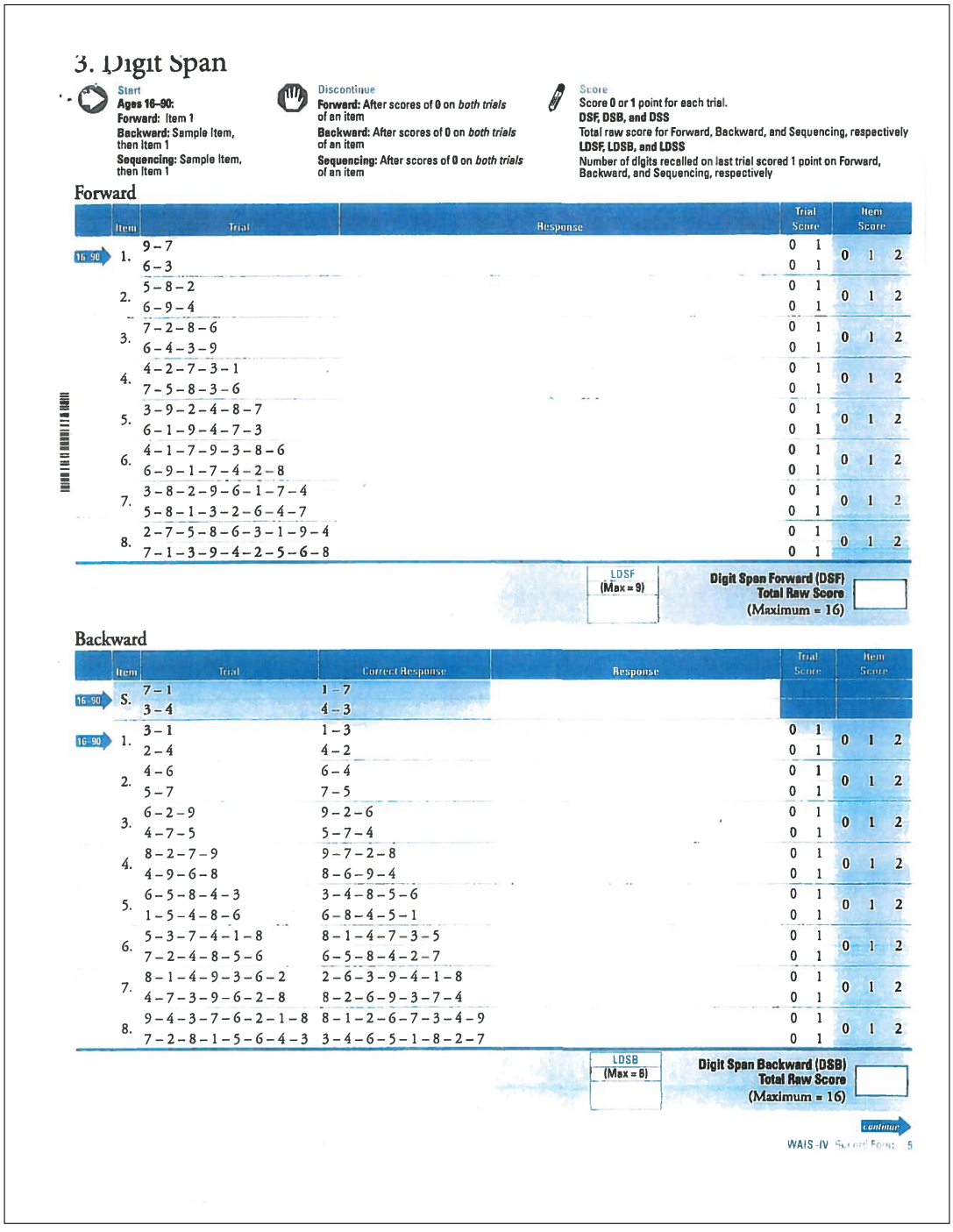


Figure D.3: Snellen Eye Acuity Test Chart, the eye chart test used to measure visual acuity.

D.6. Digit Span Memory Instructions



3. Digit Span

Start
Ages 16–90:
Forward: Item 1
Backward: Sample Item, then Item 1
Sequencing: Sample Item, then Item 1

Discontinue
Forward: After scores of 0 on *both trials* of an item
Backward: After scores of 0 on *both trials* of an item
Sequencing: After scores of 0 on *both trials* of an item

Score
Score 0 or 1 point for each trial.
DSF, DSB, and DSS
Total raw score for Forward, Backward, and Sequencing, respectively
LDSF, LDSB, and LDSS
Number of digits recalled on last trial scored 1 point on Forward, Backward, and Sequencing, respectively

Forward

Item	Trial	Response	Trial Score	Item Score
16–90 1.	9 – 7		0 1	0 1 2
	6 – 3		0 1	
2.	5 – 8 – 2		0 1	0 1 2
	6 – 9 – 4		0 1	
3.	7 – 2 – 8 – 6		0 1	0 1 2
	6 – 4 – 3 – 9		0 1	
4.	4 – 2 – 7 – 3 – 1		0 1	0 1 2
	7 – 5 – 8 – 3 – 6		0 1	
5.	3 – 9 – 2 – 4 – 8 – 7		0 1	0 1 2
	6 – 1 – 9 – 4 – 7 – 3		0 1	
6.	4 – 1 – 7 – 9 – 3 – 8 – 6		0 1	0 1 2
	6 – 9 – 1 – 7 – 4 – 2 – 8		0 1	
7.	3 – 8 – 2 – 9 – 6 – 1 – 7 – 4		0 1	0 1 2
	5 – 8 – 1 – 3 – 2 – 6 – 4 – 7		0 1	
8.	2 – 7 – 5 – 8 – 6 – 3 – 1 – 9 – 4		0 1	0 1 2
	7 – 1 – 3 – 9 – 4 – 2 – 5 – 6 – 8		0 1	

LDSF (Max = 9)

Digit Span Forward (DSF) Total Raw Score (Maximum = 16)

Backward

Item	Trial	Correct Response	Response	Trial Score	Item Score
16–90 S.	7 – 1	1 – 7			
	3 – 4	4 – 3			
16–90 1.	3 – 1	1 – 3		0 1	0 1 2
	2 – 4	4 – 2		0 1	
2.	4 – 6	6 – 4		0 1	0 1 2
	5 – 7	7 – 5		0 1	
3.	6 – 2 – 9	9 – 2 – 6		0 1	0 1 2
	4 – 7 – 5	5 – 7 – 4		0 1	
4.	8 – 2 – 7 – 9	9 – 7 – 2 – 8		0 1	0 1 2
	4 – 9 – 6 – 8	8 – 6 – 9 – 4		0 1	
5.	6 – 5 – 8 – 4 – 3	3 – 4 – 8 – 5 – 6		0 1	0 1 2
	1 – 5 – 4 – 8 – 6	6 – 8 – 4 – 5 – 1		0 1	
6.	5 – 3 – 7 – 4 – 1 – 8	8 – 1 – 4 – 7 – 3 – 5		0 1	0 1 2
	7 – 2 – 4 – 8 – 5 – 6	6 – 5 – 8 – 4 – 2 – 7		0 1	
7.	8 – 1 – 4 – 9 – 3 – 6 – 2	2 – 6 – 3 – 9 – 4 – 1 – 8		0 1	0 1 2
	4 – 7 – 3 – 9 – 6 – 2 – 8	8 – 2 – 6 – 9 – 3 – 7 – 4		0 1	
8.	9 – 4 – 3 – 7 – 6 – 2 – 1 – 8	8 – 1 – 2 – 6 – 7 – 3 – 4 – 9		0 1	0 1 2
	7 – 2 – 8 – 1 – 5 – 6 – 4 – 3	3 – 4 – 6 – 5 – 1 – 8 – 2 – 7		0 1	

LDSB (Max = 8)

Digit Span Backward (DSB) Total Raw Score (Maximum = 16)

continue

WAIS-IV Record Form 5

Figure D.4: Digit Span Memory Instructions and Test Sheet used to measure short term memory.

D.7. Trails Making Test Instructions

Trail Making Test (TMT) Parts A & B

Instructions:

Both parts of the Trail Making Test consist of 25 circles distributed over a sheet of paper. In Part A, the circles are numbered 1 – 25, and the patient should draw lines to connect the numbers in ascending order. In Part B, the circles include both numbers (1 – 13) and letters (A – L); as in Part A, the patient draws lines to connect the circles in an ascending pattern, but with the added task of alternating between the numbers and letters (i.e., 1-A-2-B-3-C, etc.). The patient should be instructed to connect the circles as quickly as possible, without lifting the pen or pencil from the paper. Time the patient as he or she connects the "trail." If the patient makes an error, point it out immediately and allow the patient to correct it. Errors affect the patient's score only in that the correction of errors is included in the completion time for the task. It is unnecessary to continue the test if the patient has not completed both parts after five minutes have elapsed.

Step 1:	Give the patient a copy of the Trail Making Test Part A worksheet and a pen or pencil.
Step 2:	Demonstrate the test to the patient using the sample sheet (Trail Making Part A – *SAMPLE*).
Step 3:	Time the patient as he or she follows the "trail" made by the numbers on the test.
Step 4:	Record the time.
Step 5:	Repeat the procedure for Trail Making Test Part B.

Scoring:

Results for both TMT A and B are reported as the number of seconds required to complete the task; therefore, higher scores reveal greater impairment.

	Average	Deficient	Rule of Thumb
Trail A	29 seconds	> 78 seconds	Most in 90 seconds
Trail B	75 seconds	> 273 seconds	Most in 3 minutes

Sources:

- Corrigan JD, Hinkeldey MS. Relationships between parts A and B of the Trail Making Test. *J Clin Psychol.* 1987;43(4):402–409.
- Gaudino EA, Geisler MW, Squires NK. Construct validity in the Trail Making Test: what makes Part B harder? *J Clin Exp Neuropsychol.* 1995;17(4):529-535.
- Lezak MD, Howieson DB, Loring DW. *Neuropsychological Assessment.* 4th ed. New York: Oxford University Press; 2004.
- Reitan RM. Validity of the Trail Making test as an indicator of organic brain damage. *Percept Mot Skills.* 1958;8:271-276.

Figure D.5: Trails Making Test Instructions for test used to measure visual processing speed.

D.8. Trails Making Test A

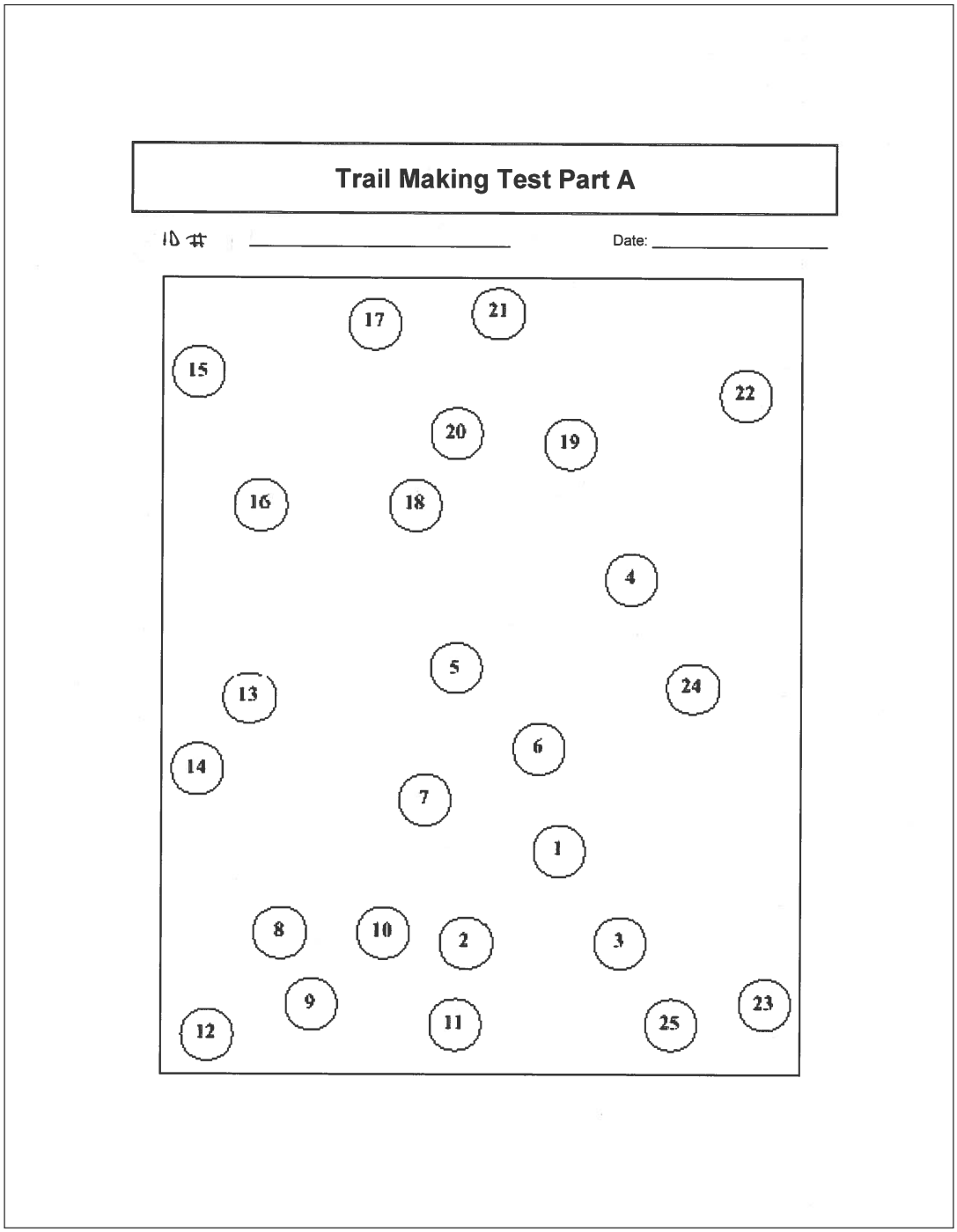


Figure D.6: Test used to measure visual processing speed.

D.9. Trails Making Test B

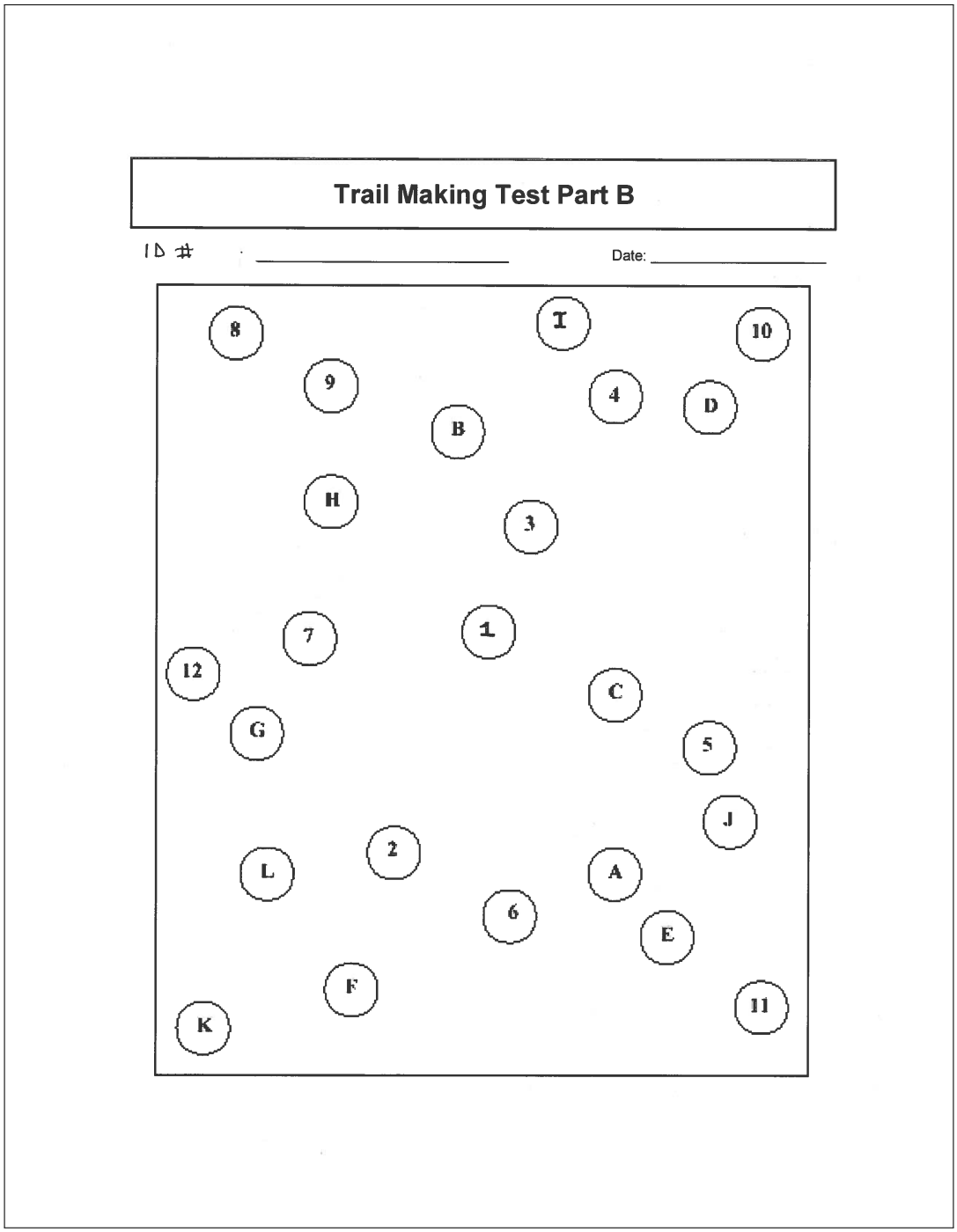


Figure D.7: Instructions for test used to measure visual processing speed.

D.10. Convoy Task Penalty Script

Table D.1: Script of scheduled *Friendly Damage* returned by route and times that route has been selected.

Selection	Route 1	Route 2	Route 3	Route 4
1	-350	0	-50	-250
2	-250	-1250	-50	0
3	0	0	0	0
4	-200	0	-50	0
5	0	0	0	0
6	-300	0	-50	0
7	0	0	0	0
8	-150	0	-50	0
9	0	0	0	0
10	0	0	0	0
11	-350	0	-50	-250
12	-250	-1250	-50	0
13	0	0	0	0
14	-200	0	-50	0
15	0	0	0	0
16	-300	0	-50	0
17	0	0	0	0
18	-150	0	-50	0
19	0	0	0	0
20	0	0	0	0
21	-350	0	-50	-250
22	-250	-1250	-50	0
23	0	0	0	0
24	-200	0	-50	0
25	0	0	0	0
26	-300	0	-50	0
27	0	0	0	0
28	-150	0	-50	0
29	0	0	0	0
30	0	0	0	0
31	-350	0	-50	-250
32	-250	-1250	-50	0
33	0	0	0	0
34	-200	0	-50	0
35	0	0	0	0
36	-300	0	-50	0

Table D.1 (continued): Script of scheduled *Friendly Damage* returned by route and times that route has been selected.

Selection	Route 1	Route 2	Route 3	Route 4
37	0	0	0	0
38	-150	0	-50	0
39	0	0	0	0
40	0	0	0	0
41	-350	0	-50	-250
42	-250	-1250	-50	0
43	0	0	0	0
44	-200	0	-50	0
45	0	0	0	0
46	-300	0	-50	0
47	0	0	0	0
48	-150	0	-50	0
49	0	0	0	0
50	0	0	0	0
51	-350	0	-50	-250
52	-250	-1250	-50	0
53	0	0	0	0
54	-200	0	-50	0
55	0	0	0	0
56	-300	0	-50	0
57	0	0	0	0
58	-150	0	-50	0
59	0	0	0	0
60	0	0	0	0
61	-350	0	-50	-250
62	-250	-1250	-50	0
63	0	0	0	0
64	-200	0	-50	0
65	0	0	0	0
66	-300	0	-50	0
67	0	0	0	0
68	-150	0	-50	0
69	0	0	0	0
70	0	0	0	0
71	-350	0	-50	-250
72	-250	-1250	-50	0
73	0	0	0	0
74	-200	0	-50	0

Table D.1 (continued): Script of scheduled *Friendly Damage* returned by route and times that route has been selected.

Selection	Route 1	Route 2	Route 3	Route 4
75	0	0	0	0
76	-300	0	-50	0
77	0	0	0	0
78	-150	0	-50	0
79	0	0	0	0
80	0	0	0	0
81	-350	0	-50	-250
82	-250	-1250	-50	0
83	0	0	0	0
84	-200	0	-50	0
85	0	0	0	0
86	-300	0	-50	0
87	0	0	0	0
88	-150	0	-50	0
89	0	0	0	0
90	0	0	0	0
91	-350	0	-50	-250
92	-250	-1250	-50	0
93	0	0	0	0
94	-200	0	-50	0
95	0	0	0	0
96	-300	0	-50	0
97	0	0	0	0
98	-150	0	-50	0
99	0	0	0	0
100	0	0	0	0
101	-350	0	-50	-250
102	-250	-1250	-50	0
103	0	0	0	0
104	-200	0	-50	0
105	0	0	0	0
106	-300	0	-50	0
107	0	0	0	0
108	-150	0	-50	0
109	0	0	0	0
110	0	0	0	0
111	-350	0	-50	-250
112	-250	-1250	-50	0

Table D.1 (continued): Script of scheduled *Friendly Damage* returned by route and times that route has been selected.

Selection	Route 1	Route 2	Route 3	Route 4
113	0	0	0	0
114	-200	0	-50	0
115	0	0	0	0
116	-300	0	-50	0
117	0	0	0	0
118	-150	0	-50	0
119	0	0	0	0
120	0	0	0	0
121	-350	0	-50	-250
122	-250	-1250	-50	0
123	0	0	0	0
124	-200	0	-50	0
125	0	0	0	0
126	-300	0	-50	0
127	0	0	0	0
128	-150	0	-50	0
129	0	0	0	0
130	0	0	0	0
131	-350	0	-50	-250
132	-250	-1250	-50	0
133	0	0	0	0
134	-200	0	-50	0
135	0	0	0	0
136	-300	0	-50	0
137	0	0	0	0
138	-150	0	-50	0
139	0	0	0	0
140	0	0	0	0
141	-350	0	-50	-250
142	-250	-1250	-50	0
143	0	0	0	0
144	-200	0	-50	0
145	0	0	0	0
146	-300	0	-50	0
147	0	0	0	0
148	-150	0	-50	0
149	0	0	0	0
150	0 -350	0	0	0

Table D.1 (continued): Script of scheduled *Friendly Damage* returned by route and times that route has been selected.

Selection	Route 1	Route 2	Route 3	Route 4
151	-250	0	-50	-250
152	0	-1250	-50	0
153	-200	0	0	0
154	0	0	-50	0
155	-300	0	0	0
156	0	0	-50	0
157	-150	0	0	0
158	0	0	-50	0
159	0	0	0	0
160	-350	0	0	0
161	-250	0	-50	-250
162	0	-1250	-50	0
163	-200	0	0	0
164	0	0	-50	0
165	-300	0	0	0
166	0	0	-50	0
167	-150	0	0	0
168	0	0	-50	0
169	0	0	0	0
170	-350	0	0	0
171	-250	0	-50	-250
172	0	-1250	-50	0
173	-200	0	0	0
174	0	0	-50	0
175	-300	0	0	0
176	0	0	-50	0
177	-150	0	0	0
178	0	0	-50	0
179	0	0	0	0
180	-350	0	0	0
181	-250	0	-50	-250
182	0	-1250	-50	0
183	-200	0	0	0
184	0	0	-50	0
185	-300	0	0	0
186	0	0	-50	0
187	-150	0	0	0
188	0	0	-50	0

Table D.1 (continued): Script of scheduled *Friendly Damage* returned by route and times that route has been selected.

Selection	**Route 1**	**Route 2**	**Route 3**	**Route 4**
189	0	0	0	0
190	-350	0	0	0
191	-250	0	-50	-250
192	0	-1250	-50	0
193	-200	0	0	0
194	0	0	-50	0
195	-300	0	0	0
196	0	0	-50	0
197	-150	0	0	0
198	0	0	-50	0
199	0	0	0	0
200	-350	0	0	0

D.11. Post Task Survey Form

Post task Survey

Subject #: Date:

Convoy Task

1. During the convoy task, how did you determine which road to select?

2. Did you use a particular strategy? If so, what was it?

3. Please rate the routes from safest (1) to most dangerous (4):

Top left road	Top right road
Bottom left road	Bottom right road

Map matching task:

1. On which map features did you sort?

2. How quickly did you realize that the sorting rule had changed? Check the response that best characterizes your overall experience.
   _____Immediately/After 1-2 trials
   _____After a few trials (3 -4 trials)
   _____After several trials (5+ trials)
   _____Did not realize sorting rule had changed

Please continue to questions on back of sheet.

Figure D.8: Post Test Survey used to gain subject feedback on tasks.

EEG:

1. How comfortable was it wearing the EEG cap?

2. Do you think it affected your performance during any of the tests? If so, how?

Additional Comments:
Are there any other comments for the study team?

Thank you for your participation!

Figure D.8 (continued): Post Test Survey used to gain subject feedback on tasks.

D.12. Decision path chart for study 2

Table D.2: Decision path chart for study 2, showing the first 20 of 96 total paths.

Path	MD1	TD1	MD2	TD2	INFO_CNT	INFO_SCORE	PATH SCORE
1	1	2	1	1	0	3	8
2	1	2	1	1	1	2	7
3	1	2	0	1	0	3	7
4	0	2	1	1	0	3	7
5	1	1	1	1	0	3	7
6	1	2	1	0	0	3	7
7	1	2	1	1	2	1	6
8	1	2	0	1	1	2	6
9	0	2	1	1	1	2	6
10	1	2	1	0	1	2	6
11	1	1	1	1	1	2	6
12	1	1	0	1	0	3	6
13	0	2	1	0	0	3	6
14	1	2	0	0	0	3	6
15	1	0	1	1	0	3	6
16	0	1	1	1	0	3	6
17	1	1	1	0	0	3	6
18	0	2	0	1	0	3	6
19	1	2	0	1	2	1	5
20	1	2	1	0	2	1	5
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
96	0	0	0	0	3	0	0

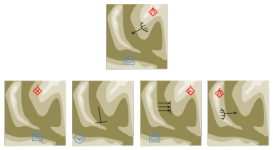
D.13. Study Synopsis for RADML Doll visit

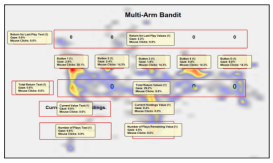
Optimal Decision Making Project
Sponsored by the Army Research Office 62626-NS
Primary Investigator: Quinn Kennedy, PhD
TRAC Analyst: MAJ Pete Nesbitt

The proposed research aims to gain insights into optimal wargaming decision making mechanisms using neurophysiological measures.

Previous research has used eyetracking and electroencephalographical (EEG) data successfully to identify when decision makers transition from exploration to exploitation, however, this research has not been done in a military decision making context (Poythress, et al, 2006; Stevens, Galloway, Berka, 2007).

The current study uses a within-person design to investigate whether brain activity and visual scan patterns via EEG and eyetracking technology can explain the development of optimal wargaming decision making; i.e., the transition from exploring to exploiting the environment. Towards this goal, we developed wargaming equivalents of the Iowa Gambling Task (IGT) (Bechara, Damasio, Damasio, & Anderson, 1994) and the Wisconsin Card Sorting Task (WCST) (Grant & Berg 1948). The IGT and WCST measure reinforcement learning and cognitive flexibility respectively, two factors necessary for optimal decision making.





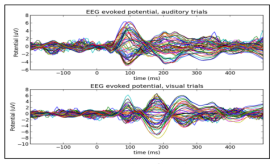


Figure 1: Pilot tests developed for this project are military versions of understood and documented decision tasks. These tests include a Map Matching task (similar to the Wisconsin Card Sorting task) and a Convoy Task (similar to the Multi-Arm Bandit or Iowa Gambling task).

In the Convoy Task (our version of the IGT), participants see four identical roads and must select which road to send a convoy (see Figure 1). Once they make a selection, they receive feedback about how many convoys were lost. Over 100 trials, participants learn to determine which roads are safe.

In the Map Matching Task (our version of the WCST), participants see five digital representations of graphical command and control plans (see Figure 1). Over several trials, they must try to correctly match the map on the top of the screen with one of the four maps at the bottom of the screen. Unbeknownst to the participant, the matching rule changes periodically. Thus, participants must not only learn the correct matching rule, but also exhibit cognitive flexibility in detecting when the rule has changed.

Our novel statistical procedures include sequential model hypothesis testing, to test the prediction that the transition from exploration to exploitation will be evidenced by the accrual of a more efficient visual scan pattern and a deduction in cognitive load (via EEG measurements).

Figure D.9: Synopsis provided to RADML Doll as an example of Naval Postgraduate School research during his visit as Commanding Officer, Navy Medicine Research.

REFERENCES

Broad agency announcement for basic and applied scientific research. Technical report, Army Research Laboratory Broad Agency Announcement, Research Triangle Park, NC, May 2012. URL `http://www.arl.army.mil/www/pages/8/Mod2_ARL_BAA_revsept13.pdf`.

LTC Jonathan Alt, LTC Jason Caldwell, MAJ Tom Deveans, MAJ James Henry, MAJ Chris Marks, MAJ Ed Masotti, and MAJ Peter Nesbitt. TRAC—Monterey FY13 research report. Draft Technical Memorandum, July 2013.

ARO. U.S. Army Research Laboratory's Army Research Office Functions. 2012. URL `http://www.arl.army.mil/www/default.cfm?page=29`.

J. Y Audibert, R. Munos, and C. Szepesvári. Tuning bandit algorithms in stochastic environments. In *Algorithmic Learning Theory*, pages 150–165, Berlin, 2007. Springer.

Jean-Yves Audibert and Sébastien Bubeck. Regret bounds and minimax policies under partial monitoring. *The Journal of Machine Learning Research*, 9999:2785–2836, 2010.

P. Auer and R. Ortner. UCB revisited: Improved regret bounds for the stochastic multi-armed bandit problem. *Periodica Mathematica Hungarica*, 61(1):55–65, 2010.

P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine learning*, 47(2):235–256, 2002. ISSN 0885-6125.

Francisco Barceló. Electrophysiological evidence of two different types of error in the wisconsin card sorting test. *Neuroreport*, 10(6):1299–1303, 1999.

Francisco Barceló and Francisco J Rubia. Non-frontal p3b-like activity evoked by the wisconsin card sorting test. *Neuroreport*, 9(4):747–751, 1998.

Antoine Bechara, Antonio R Damasio, Hanna Damasio, and Steven W Anderson. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, 50 (1):7–15, 1994.

Antoine Bechara, Hanna Damasio, Daniel Tranel, and Antonio R Damasio. The iowa gambling task and the somatic marker hypothesis: some questions and answers. *Trends in cognitive sciences*, 9(4):159–162, 2005.

Andrew H Bellenkes, Christopher D Wickens, and Arthur F Kramer. Visual scanning and pilot expertise: the role of attentional flexibility and mental model development. *Aviation, Space, and Environmental Medicine*, 1997.

Chris Berka, Daniel J Levendowski, Michelle N Lumicao, Alan Yau, Gene Davis, Vladimir T Zivkovic, Richard E Olmstead, Patrice D Tremoulet, and Patrick L Craven. Eeg correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, space, and environmental medicine*, 78(Supplement 1):B231–B244, 2007.

Dimitri Bertsekas and John Tsitsiklis. *Neuro-dynamic Programming*. Number 3 in Athena Scientific Optimization and Computation Series. Athena Scientific, Belmont, MA, 1996.

Allison AM Bielak, David F Hultsch, Esther Strauss, Stuart WS MacDonald, and Michael A Hunter. Intraindividual variability is related to cognitive change in older adults: evidence for within-person coupling. *Psychology and aging*, 25(3):575, 2010.

Avinoam Borowsky, Tal Oron-Gilad, and Yisrael Parmet. The role of driving experience in hazard perception and categorization: A traffic-scene paradigm. 2010.

John R Boyd, Chester W Richards, Franklin C Spinney, and Ginger Gholston Richards. *Patterns of conflict*. Defense and the National Interest, 2007.

Jeffrey B Brookings, Glenn F Wilson, and Carolyne R Swain. Psychophysiological responses to changes in workload during simulated air traffic control. *Biological psychology*, 42(3): 361–377, 1996.

B Rael Cahn and John Polich. Meditation states and traits: Eeg, erp, and neuroimaging studies. *Psychological bulletin*, 132(2):180, 2006.

Gwendolyn E Campbell, Christine L Belz, and Phan Luu. what was he thinking?: Using eeg data to facilitate the interpretation of performance patterns. In *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience*, pages 339–347. Springer, 2009.

Gwendolyn E Campbell, Christine L Belz, Charles PR Scott, and Phan Luu. Eeg knows best: predicting future performance problems for targeted training. In *Foundations of Augmented Cognition. Directing the Future of Adaptive Systems*, pages 131–136. Springer, 2011.

Neil Charness, Eyal M Reingold, Marc Pomplun, and Dave M Stampe. The perceptual aspect of skilled performance in chess: Evidence from eye movements. *Memory & Cognition*, 29 (8):1146–1152, 2001.

Bradley T Cowden. Modeling of helicopter pilot misperception during overland navigation. Technical report, DTIC Document, 2012.

HQ DA. *Field Manual No. FM 1-02, Operational Terms and Graphics*. Headquarters, Department of the Army, Washington, DC, September 2004. URL `https://rdl.train.army.mil/catalog/view/100.ATSC/6C01FFE5-0DF6-415F-A13C-92ED36A708CA-1303039189337/1-02/toc.htm`.

C Ervin Davis, Jessica D Hauf, D Qiang Wu, and D Erik Everhart. Brain function with complex decision making using electroencephalography. *International Journal of Psychophysiology*, 79(2):175–183, 2011.

Francesco Di Nocera, Marco Camilli, and Michela Terenzi. A random glance at the flight deck: Pilots' scanning strategies and the real-time assessment of mental workload. *Journal of Cognitive Engineering and Decision Making*, 1(3):271–285, 2007.

Mica R Endsley. Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1):32–64, 1995.

K Anders Ericsson. An introduction to cambridge handbook of expertise and expert performance: Its development, organization, and content. *The Cambridge handbook of expertise and expert performance*, pages 3–20, 2006.

Nuffield Foundation. Investigating the wisconsin card sorting test, advanced applied science: Gce a2 units. 2008. URL `http://www.nuffieldfoundation.org/sites/default/files/24_WCST_correlational.pdf`.

Ron Fricker. Introduction to statistical methods for biosurveillance. *Cambridge University Press. Draft available online at August*, 9:2010, 2010.

William J Gehring, Brian Goss, Michael GH Coles, David E Meyer, and Emanuel Donchin. A neural system for error detection and compensation. *Psychological science*, 4(6):385–390, 1993.

Sebastian Gluth, Jörg Rieskamp, and Christian Büchel. Classic eeg motor potentials track the emergence of value-based decisions. *NeuroImage*, 2013a.

Sebastian Gluth, Jörg Rieskamp, and Christian Büchel. Neural evidence for adaptive strategy selection in value-based decision-making. *Cerebral Cortex*, 2013b.

Cleotilde Gonzalez. Task workload and cognitive abilities in dynamic decision making. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 47(1):92–101, 2005.

David A Grant and Esta Berg. A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a weigl-type card-sorting problem. *Journal of experimental psychology*, 38(4):404, 1948.

Mary Hayhoe and Dana Ballard. Eye movements in natural behavior. *Trends in cognitive sciences*, 9(4):188–194, 2005.

Grit Herzmann and Tim Curran. Experts memory: an erp study of perceptual expertise effects on encoding and recognition. *Memory & cognition*, 39(3):412–432, 2011.

Clay B Holroyd and Olave E Krigolson. Reward prediction error signals associated with a modified time estimation task. *Psychophysiology*, 44(6):913–917, 2007.

Clay B Holroyd, Nick Yeung, Michael GH Coles, and Jonathan D Cohen. A mechanism for error detection in speeded response time tasks. *Journal of Experimental Psychology: General*, 134(2):163, 2005.

Clay B Holroyd, Olave E Krigolson, Robert Baker, Seung Lee, and Jessica Gibson. When is an error not a prediction error? an electrophysiological investigation. *Cognitive, Affective, & Behavioral Neuroscience*, 9(1):59–70, 2009.

Mariëtte Huizinga and Maurits W van der Molen. Age-group differences in set-switching and set-maintenance on the wisconsin card sorting task. *Developmental Neuropsychology*, 31(2):193–215, 2007.

Peter Kasarskis, Jennifer Stehwien, Joey Hickox, Anthony Aretz, and Chris Wickens. Comparison of expert and novice scan behaviors during vfr flight. In *Proceedings of the 11th International Symposium on Aviation Psychology*, 2001.

Zojiro Katoh. Saccade amplitude as a discriminator of flight types. *Aviation, space, and environmental medicine*, 68(3):205–208, 1997.

Gary Klein. Naturalistic decision making: Implications for design. Technical report, DTIC Document, 1993.

Gary Klein. Naturalistic decision making. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(3):456–460, 2008.

S Konishi, M Kawazu, I Uchida, H Kikyo, I Asakura, and Y Miyashita. Contribution of working memory to transient activation in human inferior prefrontal cortex during performance of the wisconsin card sorting test. *Cerebral Cortex*, 9(7):745–753, 1999.

Eileen Kowler. Eye movements: The past 25years. *Vision research*, 51(13):1457–1483, 2011.

Olav E Krigolson, Lara J Pierce, Clay B Holroyd, and James W Tanaka. Learning to become an expert: Reinforcement learning and the acquisition of perceptual expertise. *Journal of Cognitive Neuroscience*, 21(9):1833–1840, 2009.

Tze Leung Lai and Herbert Robbins. Asymptotically efficient adaptive allocation rules. *Advances in applied mathematics*, 6(1):4–22, 1985.

John E Laird and Robert E Wray III. Cognitive architecture requirements for achieving agi. In *Proceedings of the Third Conference on Artificial General Intelligence*, pages 1–6, 2010.

Michael F Land and Mary Hayhoe. In what ways do eye movements contribute to everyday activities? *Vision research*, 41(25):3559–3565, 2001.

Po-Ming Lee, Chia-Wei Chang, and Tzu-Chien Hsiao. Can human decisions be predicted through heart rate changes? In *Nature and Biologically Inspired Computing (NaBIC), 2010 Second World Congress on*, pages 189–193. IEEE, 2010.

Steven J Luck. An introduction to the event-related potential technique (cognitive neuroscience). 2005.

Maneuver Leader Development Strategy. Maneuver Center of Excellence, Directorate of Training and Doctrine, Fort Benning, Georgia, August 2013. URL `http://www.benning.army.mil/mcoe/maneuverconference/content/pdf/MLDS.pdf`.

Sandra P Marshall. Identifying cognitive state from eye metrics. *Aviation, space, and environmental medicine*, 78(Supplement 1):B165–B175, 2007.

Nita L Miller and Lawrence G Shattuck. A process model of situated cognition in military command and control. Technical report, DTIC Document, 2004.

Oury Monchi, Michael Petrides, Valentina Petre, Keith Worsley, and Alain Dagher. Wisconsin card sorting revisited: distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *The Journal of Neuroscience*, 21(19):7733–7741, 2001.

Chris Morey. Methodology Slide Development. TRADOC Analysis Center, September 2005.

Ronald R Mourant and Thomas H Rockwell. Strategies of visual search by novice and experienced drivers. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 14(4):325–335, 1972.

Terrence R Oakes, Diego A Pizzagalli, Andrew M Hendrick, Katherine A Horras, Christine L Larson, Heather C Abercrombie, Stacey M Schaefer, John V Koger, and Richard J Davidson. Functional coupling of simultaneous electrical and metabolic activity in the human brain. *Human brain mapping*, 21(4):257–270, 2004.

Herrmann CS Jncke L. Ott, Stier C. Musical expertise affects attention as reflected by auditory-evoked gamma-band activity in human eeg. *Neuroreport*, 24(9):445–50, 2013.

Peter P Perla. *The art of wargaming: a guide for professionals and hobbyists*. Naval Institute Press, 1990.

Daniel C Richardson and Michael J Spivey. Eye tracking: Characteristics and methods. *Encyclopedia of biomaterials and biomedical engineering*, pages 568–572, 2004.

Stuart Russell and Peter Norvig. *Artificial intelligence: A Modern Approach*. Prentice Hall, New York, NY, third edition, 2010.

Angela T Schriver, Daniel G Morrow, Christopher D Wickens, and Donald A Talleur. Expertise differences in attentional strategies related to pilot decision making. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(6):864–878, 2008.

Lisa S Scott, James W Tanaka, David L Sheinberg, and Tim Curran. The role of category learning in the acquisition and retention of perceptual expertise: A behavioral and neurophysiological study. *Brain research*, 1210:204–215, 2008.

Kimron L Shapiro and Jane E Raymond. Training of efficient oculomotor strategies enhances skill acquisition. *Acta Psychologica*, 71(1):217–242, 1989.

Jason Sherwin and Jeremy Gaston. Soldiers and marksmen under fire: monitoring performance with neural correlates of small arms fire localization. *Frontiers in human neuroscience*, 7, 2013.

Anna Skinner, Chris Berka, Lindsay Ohara-Long, and Marc Sebrechts. Impact of virtual environment fidelity on behavioral and neurophysiological response. In *The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC)*, volume 2010. NTSA, 2010.

B.F. Skinner. *The behavior of organisms*. Appleton-Century-Crofts, New York, 1938.

Michael J Spivey, DANIEL C Richardson, and Stanka A Fitneva. Thinking outside the brain: Spatial indices to visual and linguistic information. *The interface of language, vision, and action: Eye movements and the visual world*, pages 161–189, 2004.

Helen Steingroever, Ruud Wetzels, Annette Horstmann, Jane Neumann, and Eric-Jan Wagenmakers. Performance of healthy participants on the Iowa Gambling Task. *Psychological assessment*, 25(1):180, 2013.

Joseph Sullivan, Ji Hyun Yang, Michael Day, and Quinn Kennedy. Training simulation for helicopter navigation by characterizing visual scan patterns. *Aviation, Space, and Environmental Medicine*, 82(9):871–878, 2011.

Richard Sutton and Andrew Barto. *Reinforcement Learning: An Introduction*. Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA, 1998.

Csaba Szepesvari. *Algorithms for Reinforcement Learning*. Morgan & Claypool Publishers, San Francisco, 2010.

E.L. Thorndike. *Animal Intelligence*. MacMillan, New York, 1911.

Karl F Van Orden, Wendy Limbert, Scott Makeig, and Tzyy-Ping Jung. Eye activity correlates of workload during a visuospatial memory task. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 43(1):111–121, 2001.

M.M. Walsh and J.R. Anderson. Neural correlates of temporal credit assignment. In Dario Salvucci, editor, *Proceedings of the 10th International Conference on Cognitive Modeling*, pages 265–270, Philadelphia, PA, 2010. Drexell University.

Christopher John Cornish Hellaby Watkins. *Learning from delayed rewards.* PhD thesis, University of Cambridge, 1989.

Christopher Wickens and Justin Hollands. *Engineering Psychology and Human Performance*. Prentice Hall, Upper Saddle River, New Jersey, 3rd edition, 2000.

Christopher D Wickens and C Melody Carswell. Information processing. *Handbook of human factors and ergonomics*, 2:89–122, 1997.

Darrell A Worthy, Melissa J Hawthorne, and A Ross Otto. Heterogeneity of strategy use in the iowa gambling task: A comparison of win-stay/lose-shift and reinforcement learning models. *Psychonomic bulletin & review*, pages 1–8, 2012.

J Yang, Q Kennedy, J Sullivan, and M Day. Understanding cognitive processes of helicopter navigation by characterizing visual scan patterns. In *Proceedings of the 1st Conference on Pioneering Convergence Technologies, Jeju Island, Republic of Korea, February*, volume 14, page 16, 2011.

G. Zacharias, J. MacMillan, and S. Editors Van Hemel. *Behavioral Modeling and Simulation: From Individuals to Societies*. National Academies Press, Washington, D.C., 2008.