TRAC-M-TM-15-011
February 2015

Understanding Optimal Decision-making in War-gaming II

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

This study cost the Department of Defense approximately \$126,901.81 expended by TRAC in Fiscal Years 14-15.
Prepared on 20150206
TRAC Project Code # 060105.

TRAC-M-TM-15-011
February 2015

Understanding Optimal Decision-making in War-gaming II

MAJ Cardy Moten III

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

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)*
02-06-2015	Technical Memorandum	

4. TITLE AND SUBTITLE
Understanding Optimal Decision-making in War-gaming II

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
MAJ Cardy Moten III

5d. PROJECT NUMBER
060105

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-TM-15-011

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Army Research Office, Research Triangle Park, NC

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 purpose of this memorandum is to provide documentation of research for the Army Research Office (ARO) by the TRADOC Analysis Center, Monterey (TRAC-MTRY). The focus of the research is to develop a model that represents the relationship between neurophysiological metrics and optimal decision making. The research team modified two well-known psychological tests for a military context. The Iowa Gambling Task (IGT) was modified to assess reinforcement learning and the Wisconsin Card Sorting Test (WCST) was modified to assess cognitive flexibility. The tests were administered to 34 military officers across all services. Based on the results of the IGT and WCST, the research team also developed the Cognitive Alignment With Performance Targeted Training Intervention Model (CAPTTIM) to assess the relationship between a subject's cognitive state and their observed performance. Through analyzing reinforcement learning and cognitive flexibility, the CAPTTIM can be used to provide a real-time notification of when a training intervention is required and the type of training intervention necessary. The results indicate that the modified versions of the IGT and WCST along with the CAPTTIM can be as an objective assessment tool in conjunction with other virtual and live military decision making training.

15. SUBJECT TERMS

Training intervention, cognitive state, regret, neurophysiological measures, Wisconsin Card Sorting Test, Iowa Gambling Task

16. SECURITY CLASSIFICATION OF: Unclassified			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** MAJ Cardy Moten III
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**	UL	77	**19b. TELEPHONE NUMBER** *(include area code)*
Unclassified	Unclassified	Unclassified			831-656-2452

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

This page intentionally left blank.

Table of Contents

This page intentionally left blank.

1. **Purpose.** The purpose of this memorandum is to provide documentation of research for the Army Research Office (ARO) by the TRADOC Analysis Center, Monterey (TRAC-MTRY). The focus of the research is to develop a model that represents the relationship between neurophysiological metrics and optimal decision making.

2. **Background.** The U.S. Army published its operating concept in October of 2014. The purpose of this concept is to describe how the Army will operate at the strategic, operational, and tactical level without knowing much about the future environment, location, and enemy.[1]In order to accomplish this objective, the training for Army officers has to focus on adaptive decision making through realistic training in actual and virtual environments.[2] Currently, the metrics used in training to evaluate the decision making of officers is subjective, and little is known about how military officers make optimal decisions. A potential solution to this problem is to combine human-in-the-loop wargames with behavioral and neurophysiological measures.

3. **Methodology.**The research team modified two well-known psychological tests for a military context. The Iowa Gambling Task (IGT) was modified to assess reinforcement learning.[3] The Wisconsin Card Sorting Test (WCST) was modified to assess cognitive flexibility.[4] The tests were administered to 34 military officers across all services. Kennedy et al. discuss in detail the modification of these tests and the results of their research.[5] Based on the results of the IGT and WCST, the research team also developed the Cognitive Alignment With Performance Targeted Training Intervention Model (CAPTTIM) to assess the relationship between a subject's cognitive state and their observed performance. Through analyzing reinforcement learning and cognitive flexibility, the CAPTTIM can be used to provide a real-time notification of when a training intervention is required and the type of training intervention necessary (See Appendix C). This is done through using quantitative statistical methods to determine if a decision maker is in an exploration versus exploitation cognitive state and if they are yielding the optimal decision performance while in that particular state. In this research that decision performance metric is the amount of regret, which we define as the difference between the maximum benefit that could be received at a particular state minus the value of the benefit actually obtained. An exploration cognitive state indicates the subject is more of a naïve decision maker and needs more information on their environment.[6] An exploitation cognitive state indicates the subject is more

[1]U.S., Department of the Army Training and Doctrine Command. *TRADOC Pamphlet 525-3-1, The U.S. Army Operating Concept: Win In a Complex World*. Washington DC: Government Printing Office, October 2014.

[2]Ibid.

[3]Antoine Bechara et al. "Insensitivity to future consequences following damage to human prefrontal cortex". In: *Cognition* 50.1 (1994), pp. 7–15.

[4]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." In: *Journal of experimental psychology* 38.4 (1948), p. 404.

[5]Quinn Kennedy, Peter Nesbitt, and Jon Alt. "Assessment of Cognitive Components of Decision Making with Military Versions of the IGT and WCST". in: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 58. 1. SAGE Publications. 2014, pp. 300–304.

[6]Ibid.

experienced and has figured out the optimal alternative and does not consider any other sub-optimal alternative from that point on.[7]

4. **Progress** The sponsor was briefed on the results of the modifications to the IGT and the WCST and the development of CAPTTIM on 09 October 2015. Kennedy et al. presented their findings on the modifications to the IGT and WCST at 2014 the Human Factors Ergonomics Society annual meeting. The research team is also drafting manuscripts to submit to the Military Psychology Journal discussing their findings on the modification of the IGT and WCST. Other manuscripts being drafted during FY 15 also include:

   - Exploratory analysis of the modified IGT and WCST data.
   - Using reinforcement learning algorithms to model human decision making on the modified IGT and WCST.
   - Compare the performance of the reinforcement learning algorithms to model human decision making on the modified IGT and WCST.
   - Analyze the role of working memory and visual processing speed on military decision making.

5. **Results.** Kennedy et al. found that the tested subjects scored on average with a normed population on the IGT and below average with a normed population of the WCST.[8] The results indicate that both tests are suitable assessment tools that could be used in conjunction with other virtual and live military decision making training. The analysis of the CAPTTIM model showed that by examining the relationship between a subject's cognitive state and their optimal decision performance. If a subject is in an exploration cognitive state and has low regret, then they are risk averse and require a training intervention. Conversely, if a subject is in an exploitation cognitive state and has high regret, then they are making too many risky decisions and require a training intervention. Future work on this model seeks to to find the threshold that will define the optimal balance between an exploitation cognitive state and low regret.

[7] Ibid.
[8] Ibid.

Appendix A
Study Plan

Problem Statement

To investigate 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.

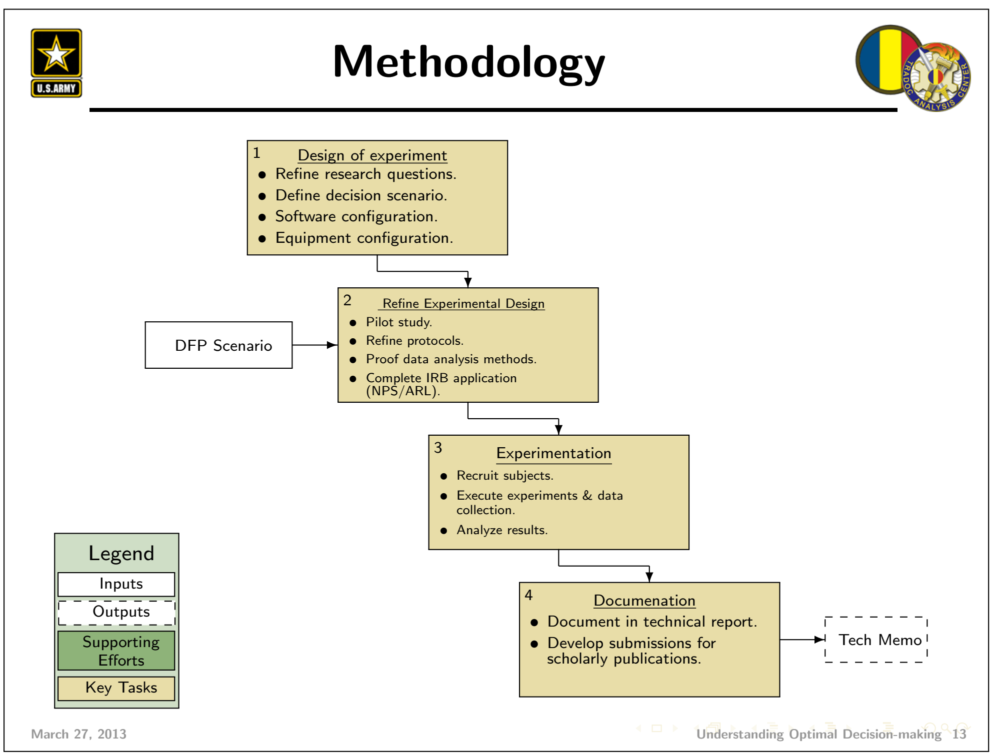
Project Team

Sponsor Agency:	Dr. Virginia Pasour Biomathematics Program U.S. Army Research Office, Research Triangle Park, NC virginia.b.pasour.civ@mail.mil
TRAC Lead:	Peter A. Nesbitt MAJ, AR/FA49 TRADOC Analysis Center - Monterey peter.a.nesbitt.mil@mail.mil
Primary Investigator:	Dr. Quinn Kennedy Operations Research Department Naval Postgraduate School, Monterey, CA mqkenned@nps.edu
NPS Faculty:	LTC Jonathan Alt Operations Research Department Naval Postgraduate School, Monterey, CA jkalt@nps.edu
	Dr. Ronald D. Fricker Operations Research Department Naval Postgraduate School, Monterey, CA rdfricke@nps.edu

Constraints, Limitations, & Assumptions

- **Constraints**
  - The total budget for this phase of the project is \$100K.
  - Phase II must be complete no later than 30 December 2014.
- **Limitations**
  - Will limit initial experimentation to discrete decision situations or with limited exposure of sequential tasks.
  - Subjects limited to those officer students available at NPS.
- **Assumptions**
  - Results of experimentation with available subject pool will be sufficient to provide insight into study issues.

Methodology



Timeline

APR 14	Submit IGT and WCST modification paper to the Human Factors and Ergonomics Society (HFES)
OCT 14	ODM II IPR
OCT 14	Present findings at the HFES annual meeting.
DEC 14	CAPTIIM Tech Report complete.

This page intentionally left blank.

Appendix B
Progress Report

The final IPR, presented to the sponsor on 09 October 2015, for this phase of the project is on the following pages.

Understanding Optimal Decision Making in Wargaming

TRAC Project 060105

Research Progress Brief to Dr. Pasour

9 October 2014

Purpose & Agenda

The purpose of this brief is to document progress of the Optimal Decision Making Project.

Agenda

- Research Overview
- ODM II Accomplished Goals
- ODM III Goals
- Follow-on Research

Research Overview

1. Problem Definition. ODM I
- Determined cognitive abilities for military decision making.
- Created simple wargaming tasks.
- Acquired neurophysiological equipment.

2. Conduct Wargaming Study. ODM II
- Collect data on 34 Officers.
- Conduct initial analysis.

→
- Completed Student Thesis.
- HFES Conference Paper.

3. Refine Methods. ODM III
- Refine statistical methods.
- Document results.
- Share deliverable products.

→
- Journal Articles.
- Collaboration with Veteran's Affairs.
- Address advanced decision making: Make Goal Task.
- Determine if results are applicable for training intervention.

Legend
- Input
- Output
- Key Tasks

ODM II Accomplished Goals

Study 1

- 34 Officers completed:
  - Map Task.
  - Convoy Task.
  - Covariate Measures.

- Synchronization of decision and EEG data

- Submitted to Professional Publications.
  - Conference Paper Accepted, Presenting at HFES Conference.
  - Manuscript submitted to Military Psychology.

- Collaboration of Results.
  - Transfer of measure to Veterans Affairs for TBI population.

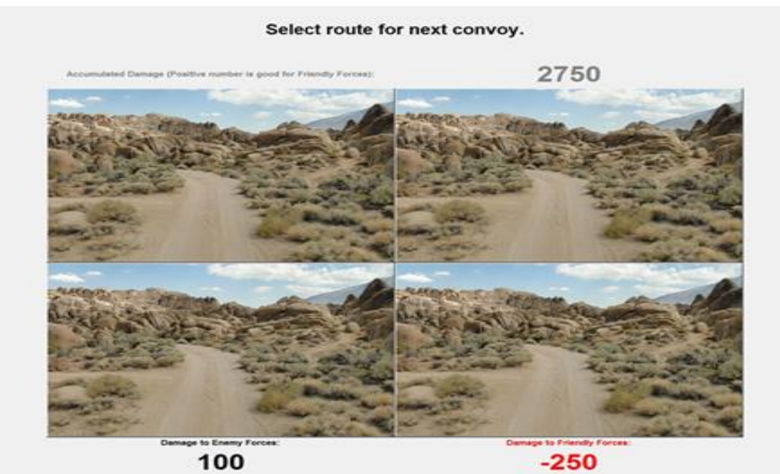
Study 2

- 30 Officers Completed.
  - Tactical Decision-making: Live versus Autonomous.

- Completed Thesis.

ODM II, Study 1, Convoy Task

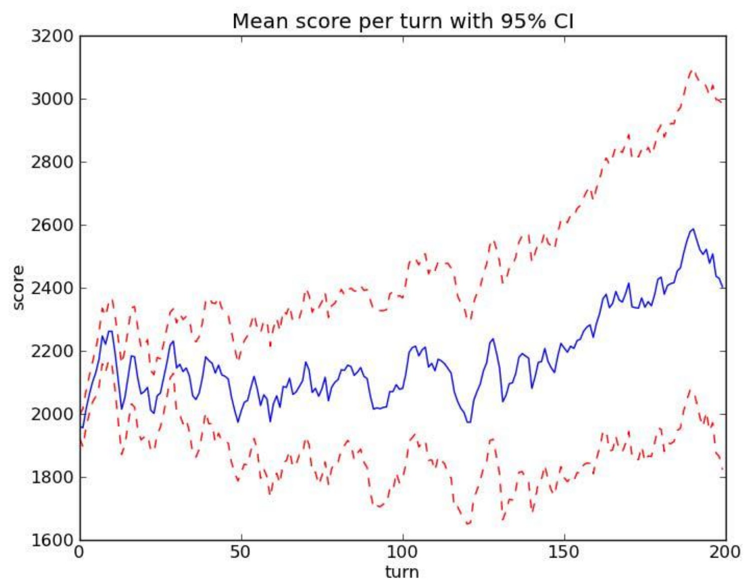
Screenshot of Task.



Total Damage score by trial indicates:

- More than 100 trials was necessary (compared to original IGT).
- Large amounts of individual variability.

Mean Total Damage Score (all participants).



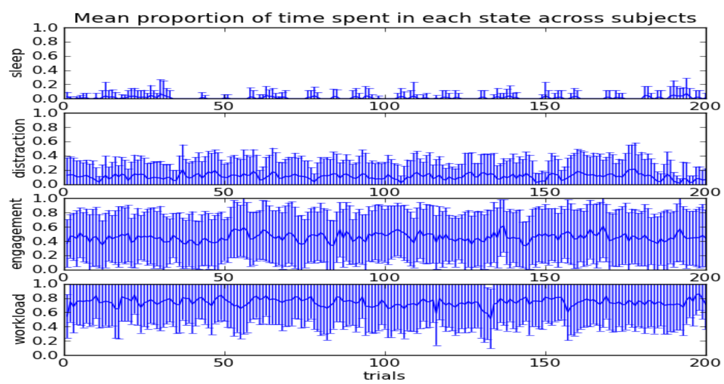
Expected Outcome.

	Route 1	Route 2	Route 3	Route 4
Min.	-250	-1150	0	-200
Median	25	100	25	50
Mean	-25	-25	25	25
Max.	100	100	50	50

The convoy task proves valid and sufficiently difficult.

ODM II, Preliminary EEG results

We can examine EEG activity on a trial by trial basis for each subject.



The convoy task was sufficiently challenging and engaging.

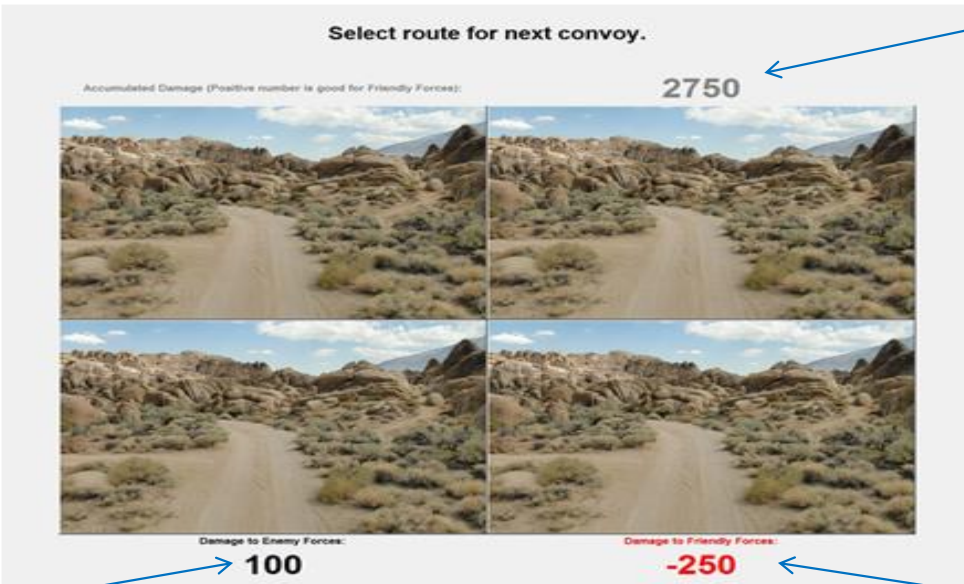
- Subjects were engaged almost 50% of time.
- Subjects experienced cognitive workload 75% of time.

	Sleepiness (%)	Distracted (%)	Engaged (%)	Workload (%)
Mean dwell time (sd)	.012 (.016)	.109 (.146)	.464 (.218)	.745 (.230)

The success of monitoring cognitive state fluctuations allows analysis of decision making performance with EEG to guide training interventions.

ODM II, Preliminary Eye Tracking Results

Total Damage



Enemy Damage

Friendly Damage

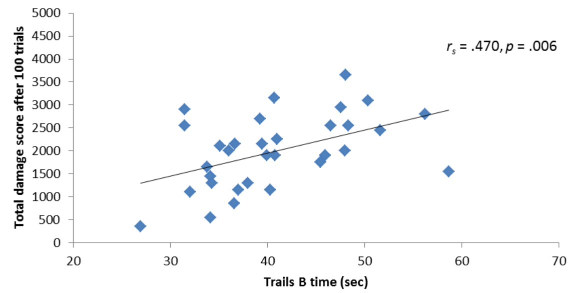
		Total Damage	Friendly Damage	Enemy Damage	Routes
Gaze time (%)	Mean (sd)	5.49 (12.47)	16.73 (14.87)	6.55 (6.40)	71.23 (19.86)
Mean dwell time (sec)	Mean (sd)	.171 (.240)	.456 (.269)	.435 (.844)	1.486 (1.195)
Median dwell time (sec)	Mean (sd)	.095 (.134)	.320 (.215)	.201 (.124)	.671 (.330)

Participants relied on friendly damage information in making their decisions (suggesting risk strategies).

ODM II, preliminary statistical modeling: covariates and cluster analyses

Covariance Measures.

Trails B, a processing speed measure, explains some of the variability in decision performance, particularly in the first 100 trials.



Cluster Analysis.

Behavior

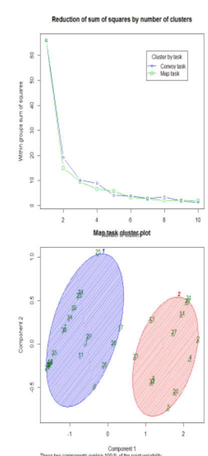
Decision Data

Performance Criteria

Final Damage

Adv. Sel. Bias

Cluster Analysis



Grouping by Performance

Performance Variable	Group 1	Group 2
First 100 Trials	**mean (sd)**	**mean (sd)**
No. trials w/ friendly damage	25.67 (9.0)	23.9 (4.7)
No. trials w/ heavy friendly damage	2.92 (1.4)	4 (1.2)
Trials 101-200		
No. trials w/ friendly damage	29.17 (10.8)	25.27 (4.5)
No. trials w/ heavy friendly damage	1.42 (1)	3.95 (1.3)
All 200 Trials		
No. trials w/ friendly damage	54.8 (15.0)	49.14 (7.9)
No. trials w/ heavy friendly damage	4.3 (1.8)	7.95 (2.0)

Group 1 Sel. rate

7 %	24 %
37 %	32 %

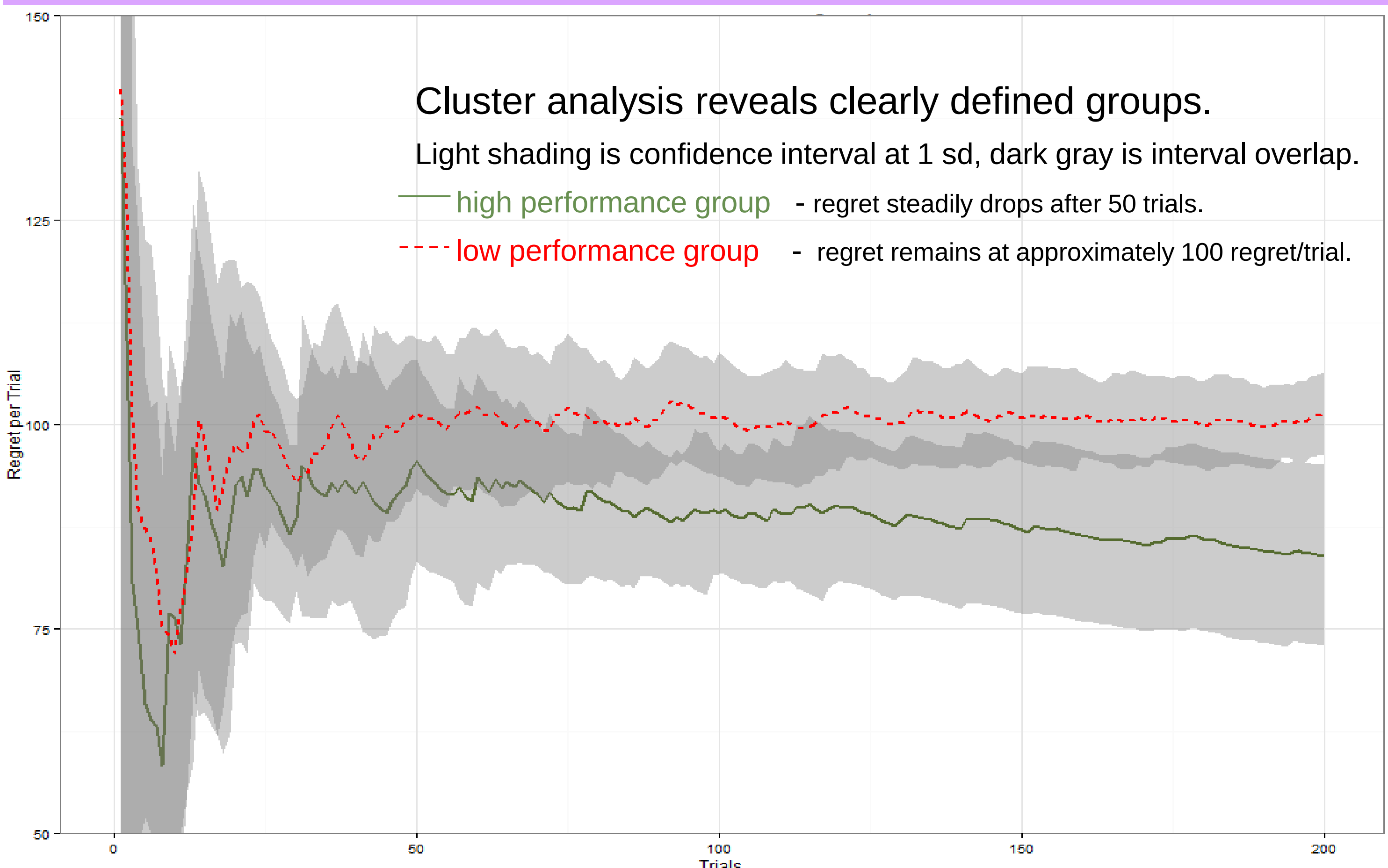
Group 2 Sel. rate

17 %	41 %
19 %	24 %

Cluster Analysis successfully distinguished between high and low performers.

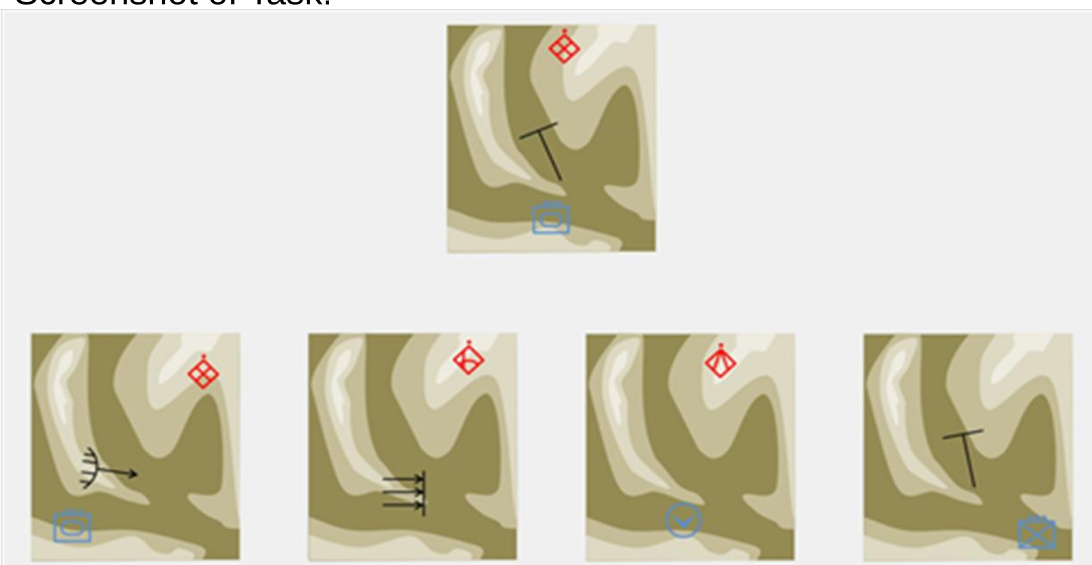
ODM II, preliminary statistical analyses: Regret combined with Clustering

Regret is the difference of a single trial outcome and ideal decision given perfect knowledge.



ODM II, Study 1, Map Task

Screenshot of Task.



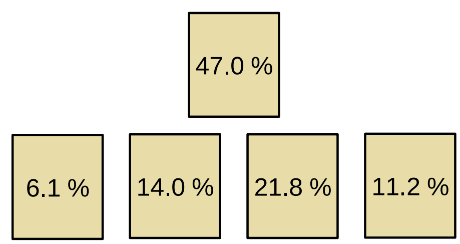
Partial Graphic Key

	friendly graphics
Level 0	no friendly graphic
Level 1	friendly armor platoon
Level 2	friendly aerial vehicle
Level 3	friendly infantry platoon

Descriptive Statistics on Map Task.

Variable	Mean (sd)	median	range
Number of trials completed	119.35 (16.52)	128	76 -128
Perseverative responses	11.82 (11.12)	9	0-37
Non perseverative errors	41.85 (22.52)	38	8-81
Number of trials to complete first rule	42.9 (28.95)	34	14-121
Number of rules achieved	3.21 (1.94)	4	0-5
Failure to maintain set	2.32 (1.49)	2	0-5

Mean Percent time

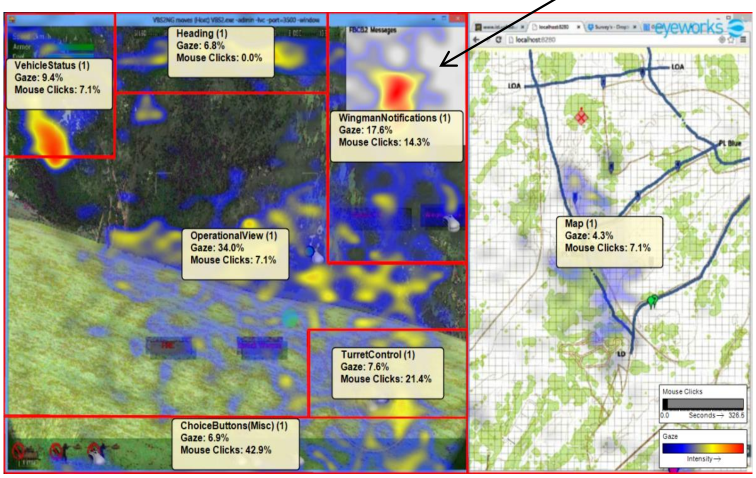


The map task proves valid and sufficiently difficult.

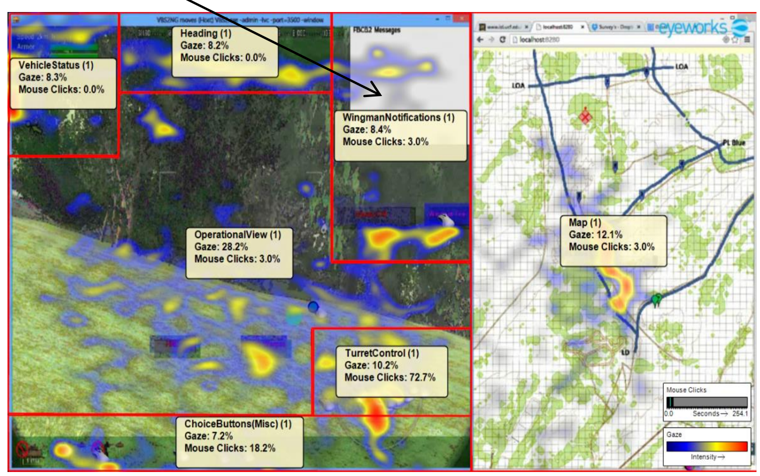
ODM II Study 2: Tactical decision making with live or automated wingman

Is performance different when tactical leaders rely on an autonomous wingman or a live wingman?

Wingman notification region of interest (ROI)



Automated wingman screenshot



Live wingman screenshot

Subjects with autonomous wingman spent significantly more time looking on the wingman notification ROI than the subject with a live wingman.

Visual scan patterns can indicate the amount of trust in wingman.

ODM III Goals

Study 1

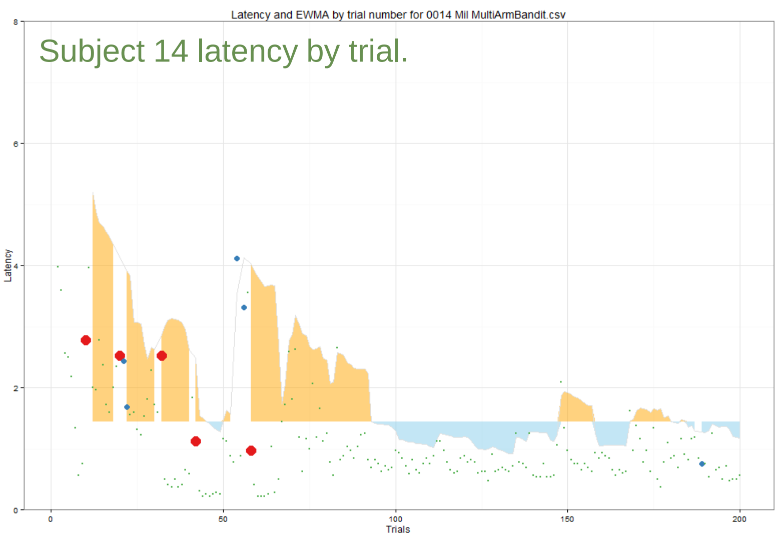
- Determine feasibility of training intervention model based off of exploration/exploitation statistical modeling.
- Submit results to peer-reviewed journal articles and conference papers.

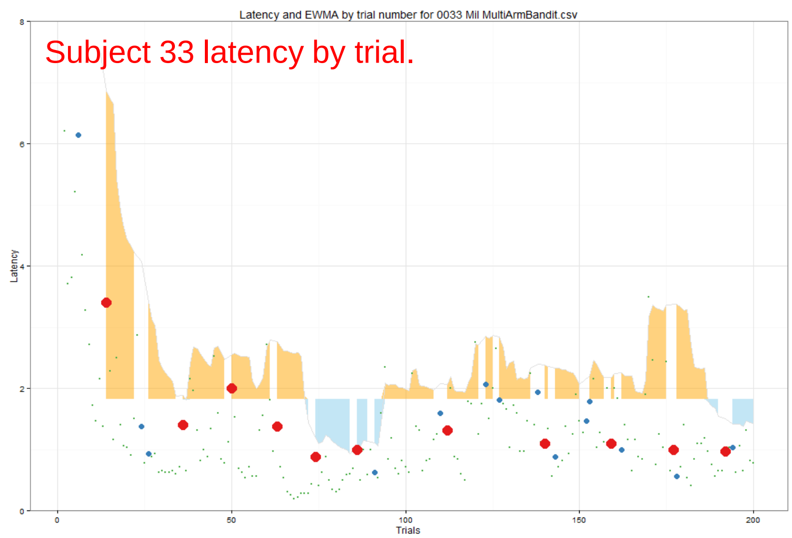
Study 2

- Determine if there are neurophysiological differences between subjects with live vs autonomous wingman.
- Submit results regarding neurophysiological differences to peer-reviewed journal article or conference paper.

ODM III, Exploration vs. Exploitation

Use of sequential sample variances in latency times to determine exploration and cognitive states.





Exploration

Exploitation

Subject 14

- Ideal transition from exploration to exploitation.

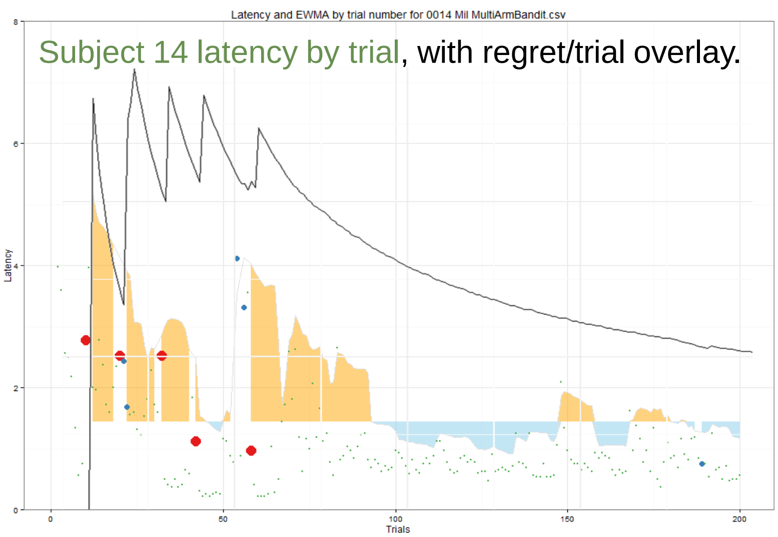
Subject 33

- Nonoptimal pattern of being almost exclusively in exploration mode throughout the task.

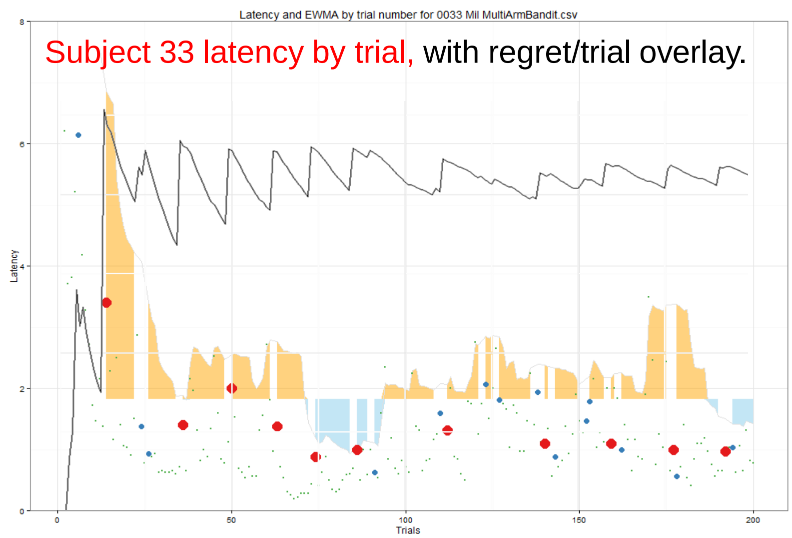
ODM III, Exploration vs. Exploitation

Use of sequential sample variances in latency times to determine exploration and cognitive states.

Subject 14 latency by trial, with regret/trial overlay.



Subject 33 latency by trial, with regret/trial overlay.



Exploration | Exploitation

Subject 14

- Ideal transition from exploration to exploitation.
- Regret is high during exploration.
- Regret decreases during exploitation mode.

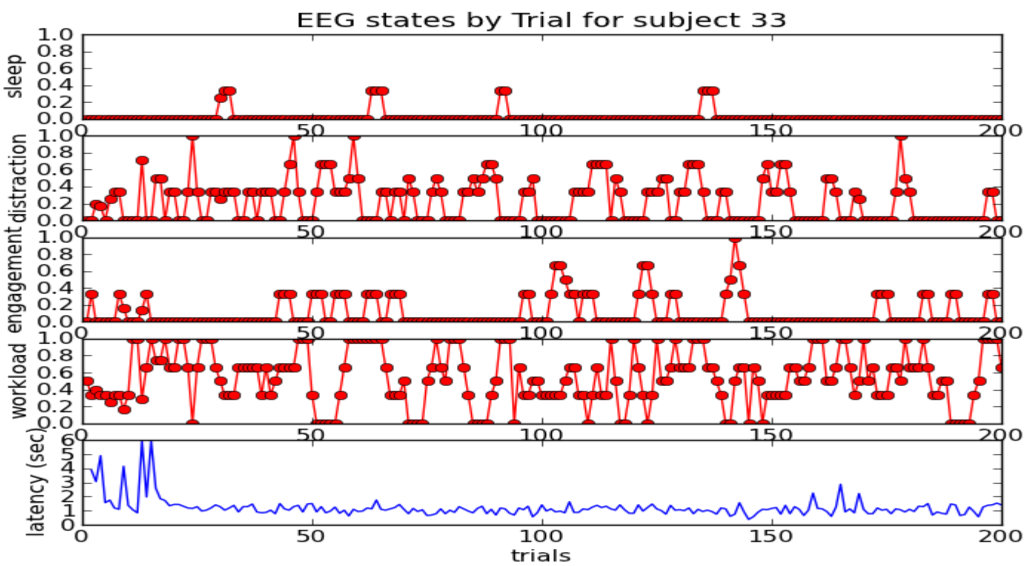
Subject 33

- Nonoptimal pattern of being almost exclusively in exploration mode throughout the task.
- Regret is high during exploration.
- No consistant decrease in regret.

Trial by trial regret is consistent with subjects' exploration and exploitation mode.

ODM III, Insights from Initial Results

We are able to align decision, covariate, EEG and eye tracking data.



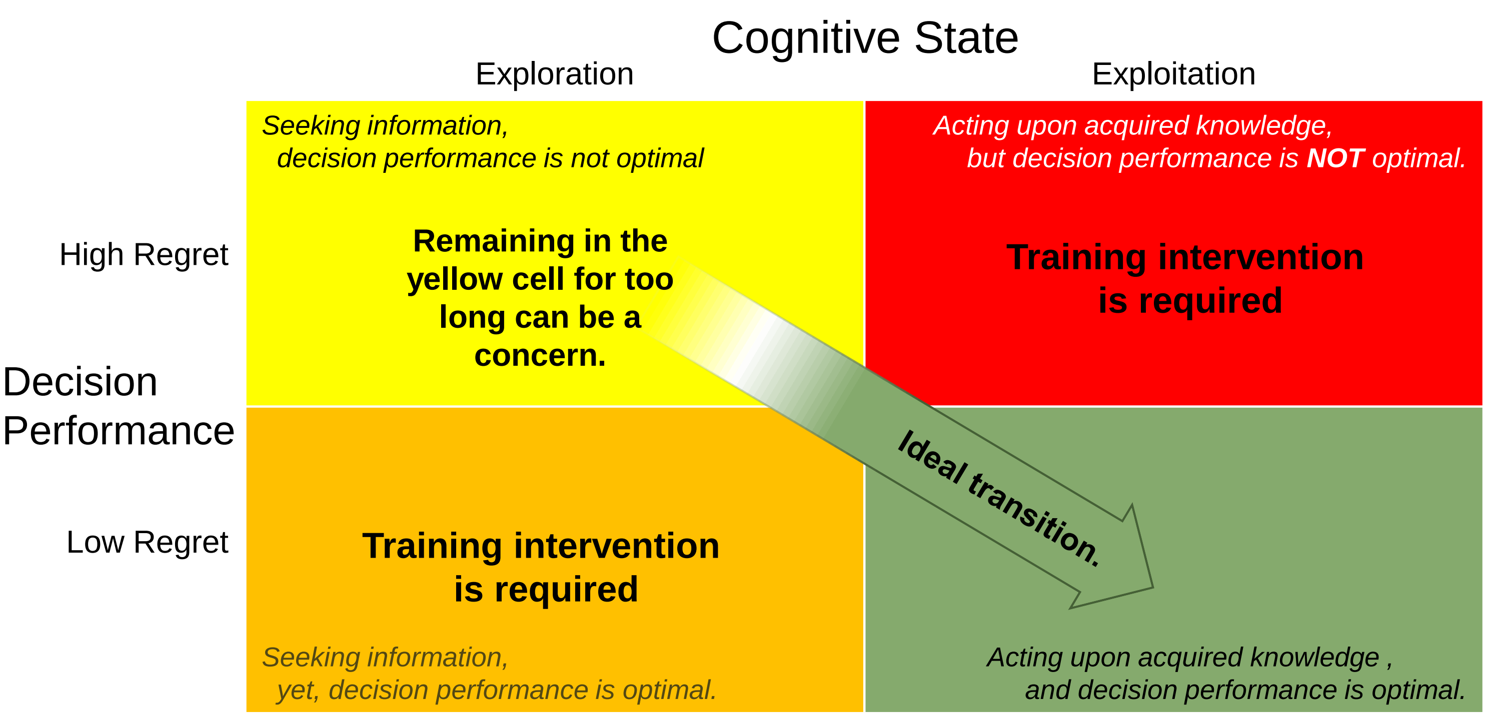
Subject 33

Experienced high frequencies of cognitive workload and distraction; providing some insight into why they were predominantly in exploration mode and had poor decision performance.

Combined data allows insight to correlations of cognitive state and performance.

ODM III, Insights from Initial Results

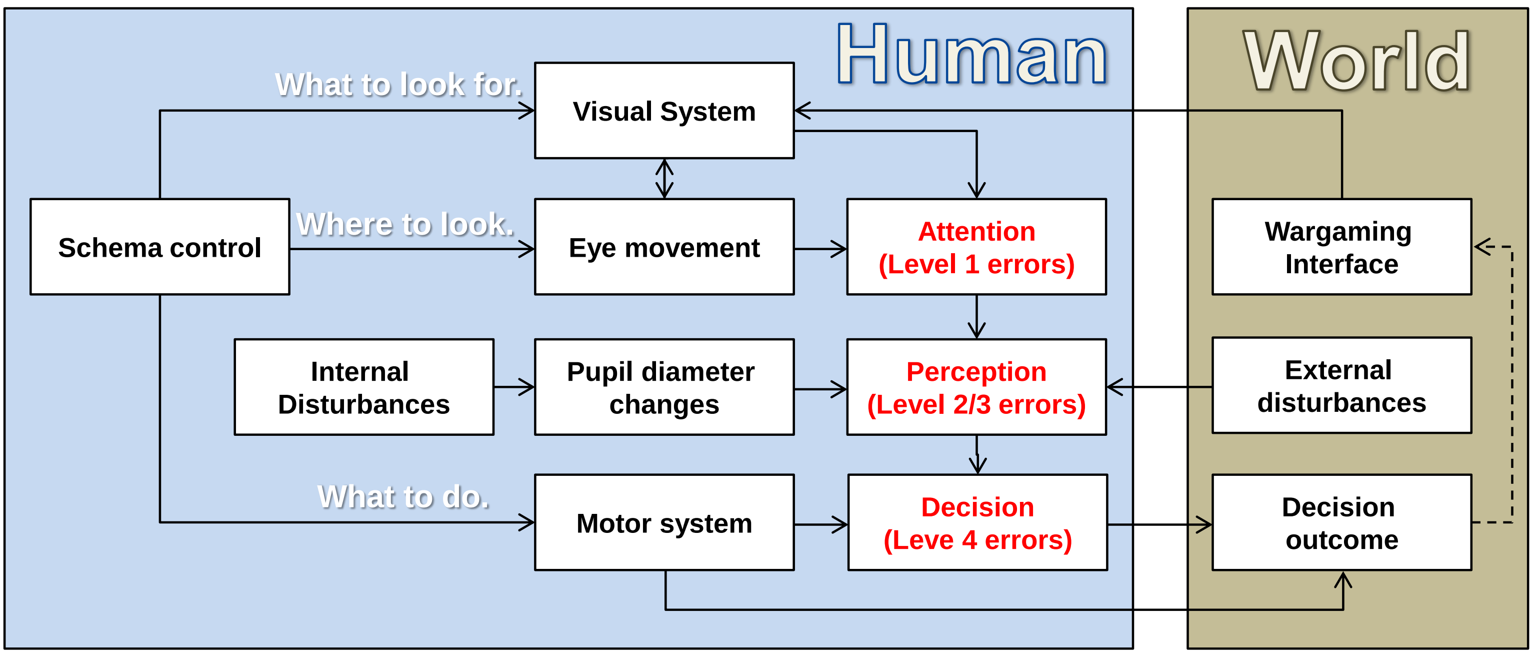
CAPTIMM: Cognitive Alignment with Performance Targeted Training Intervention Model (Kennedy et al, in preparation)



Simple behavioral variables measured and recorded in real time can be used for a near immediate training intervention.

Model of Non-Optimal Decision Making

When misalignment between cognitive state and decision performance occurs, we can use our model of non optimal decision making to understand ***WHY***.



Follow-on Research

Use knowledge acquired from ODM to seek follow-on research supported by:

Army Research Laboratories.

Office of Naval Research.

War Related Illness and Injury Study Center.

Research Team

Dr. Quinn Kennedy	Principal Investigator
MAJ Peter Nesbitt	TRAC Project Lead.

LTC Jonathan Alt
Dr. Ji Hyun Yang
Dr. Ronald D. Fricker
Dr. Jeff Appleget
Mr. Jesse Huston
MAJ Scott Patton
Mr. Lee Whitaker

Discussion and Questions

POC: MAJ Peter Nesbitt

peter.a.nesbitt.mil@mail.mil

831-656-7575

Dr. Quinn Kennedy

mqkenned@nps.edu

931-656-2618

This page intentionally left blank.

Appendix C
Cognitive Alignment With Performance Targeted Training Intervention Model Tech Report

The following pages contain the technical report for the CAPTTIM developed by Kennedy et al. Distribution is unlimited.

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

COGNITIVE ALIGNMENT WITH PERFORMANCE TARGETED TRAINING INTERVENTION MODEL: CAPTTIM

by

Quinn Kennedy, Ph.D.; MAJ Peter Nesbitt, USA; LTC Jon Alt, USA; and Ronald D. Fricker, Jr., Ph.D.

January 2015

Approved for public release; distribution is unlimited.

Prepared for: Army Research Office
PO Box 12211
Research Triangle Park, NC 27709-2211

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)***
31-01-2015	Technical Report	September 2014 - December 2014

4. TITLE AND SUBTITLE
Cognitive Alignment with Performance Targeted Training Intervention Model: CAPTTIM

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Quinn Kennedy, Ph.D.; MAJ Peter Nesbitt, USA; LTC Jon Alt, USA; and Ronald D. Fricker, Jr., Ph.D.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) AND ADDRESS(ES)
Operations Research Department
Naval Postgraduate School
Monterey, CA 93943

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Army Research Office (ARO) PO Box 12211, Research Triangle Park, NC, 27709-2211

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
The views expressed in this report are those of the author(s) and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

14. ABSTRACT
In this technical report, we propose that the use of two simple behavioral measures, in conjunction with neurophysiological measures, can be used to create a training intervention that has the potential to provide: (1) real-time notification as to when a training intervention is needed and (2) real-time information as to the type of training intervention that should be employed. The Cognitive Alignment with Performance Targeted Training Intervention Model (CAPTTIM) determines if a trainee's cognitive state is aligned or misaligned with actual performance. When misalignment occurs, it indicates that a training intervention is needed. Neurophysiological markers, as captured by eyetracking and electroencephalography (EEG), can assist in determining why misalignment between cognitive state and performance occurred, leading to more effective and targeted training intervention. Because all measures are captured continuously in real time, this model has the potential to increase training efficiency and effectiveness in a variety of training domains. The model is illustrated with two case studies.

15. SUBJECT TERMS

Training intervention, cognitive state, regret, neurophysiological measures

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			
Unclassified	Unclassified	Unclassified	UU	33	Quinn Kennedy
					19b. TELEPHONE NUMBER ***(include area code)*** (831) 656-2618

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

THIS PAGE INTENTIONALLY LEFT BLANK

NAVAL POSTGRADUATE SCHOOL
Monterey, California 93943-5000

The report entitled "Cognitive Alignment with Performance Targeted Training Intervention Model: CAPTTIM" was prepared for "Army Research Office" and funded by "Army Research Office".

Further distribution of all or part of this report is authorized.

This report was prepared by:

Quinn Kennedy
Senior Lecturer of Operations Research

Peter Nesbitt
MAJ, USA

Jon Alt
Assistant Professor of Operations Research

Ronald D. Fricker, Jr.
Professor of Operations Research

Reviewed by:

Johannes O. Royset
Associate Chairman for Research
Department of Operations Research

Robert F. Dell
Chairman
Department of Operations Research

Released by:

Jeffrey D. Paduan
Vice President and
Dean of Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

In this technical report, we propose that the use of two simple behavioral measures, in conjunction with neurophysiological measures, can be used to create a training intervention that has the potential to provide: (1) real-time notification as to when a training intervention is needed, and (2) real-time information as to the type of training intervention that should be employed. The Cognitive Alignment with Performance Targeted Training Intervention Model (CAPTTIM) determines if a trainee's cognitive state is aligned or misaligned with actual performance. When misalignment occurs, it indicates that a training intervention is needed. Neurophysiological markers as captured by eyetracking and electroencephalography (EEG) can assist in determining why misalignment between cognitive state and performance occurred, leading to more effective and targeted training intervention. Because all measures are captured continuously in real time, this model has the potential to increase training efficiency and effectiveness in a variety of training domains. The model is illustrated with two case studies.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

A. MOTIVATION

As the Army focuses on enhancing leader development and decision making to improve the effectiveness of combat forces, the importance of understanding how to effectively train decision makers and how experienced decision makers arrive at optimal or near-optimal decisions has increased. Currently, there is little understanding of how military decision makers arrive at optimal decisions and the measurement of decision-making performance lacks objectivity. The combined use of behavioral and neurophysiological measures in human-in-the-loop wargames has the potential to fill this knowledge gap and provide more objective measures of decision-making performance.

B. 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. We focused on the development of optimal decision making when all subjects begin as naïve decision makers. Specifically, we attempted 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.

C. ARMY RELEVANCY AND MILITARY APPLICATION AREAS

Objectively defining, measuring, and developing a means to assess military optimal decision making has the 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, with it, advance the understanding and modeling of decision-making processes to more deeply comprehend the fundamentals of Soldier cognition.

D. SUMMARY OF CURRENT STATUS

We developed a wargame and conducted a study that demonstrated that it successfully elicits cognitive flexibility and reinforcement learning. Based on quantitative measures of exploration and exploitation, we developed the Cognitive Alignment with Performance - Targeted Training Intervention Model (CAPTTIM). Based on real-time measures of a trainee's cognitive state and their actual performance, the model proposes a method for identifying (1) whether or not a trainee’s cognitive state is aligned or misaligned with actual performance, and (2) possible reasons as to why cognitive misalignment is occurring. We find that the combination of knowledge of cognitive state and actual decision performance gives insight into the optimality of trainees’ decisions.

I. INTRODUCTION

A. OVERVIEW

As the U.S. Army focuses on enhancing leader development and decision making to improve the effectiveness of its combat forces, the importance of understanding how to effectively train decision makers and how experienced decision makers arrive at optimal or near-optimal decisions has increased (Lopez, 2011). In order to understand how to effectively train decision makers to make optimal decisions, there are at least two components that need to be understood and quantitatively characterized. One such component is the cognitive state of the decision maker trainee: do they think they need to learn more about the environment before they can make good decisions or do they think they are making good decisions? In our work, we call this first cognitive state *exploration*: needing to learn about one's environment and actively seeking and responding to information in the environment. We refer to the latter state as *exploitation*: thinking that you have figured out the task and acting on that knowledge.

A second component of understanding optimal military decision making is having an objective measure of a trainee's actual decision performance. Ideally, this measure should provide, at any point during the task, information as to how close a trainee is to making optimal decisions. It is important to note that both components, knowledge of the decision maker's cognitive state and a measure of their actual decision performance are necessary to truly understand optimal military decision making. In the process of operationalizing the definitions of exploration and exploitation, and determining an objective measure of decision performance, we developed the Cognitive Alignment with Performance-Targeted Training Intervention Model (CAPTTIM). The purpose of this paper is to describe the model and then to illustrate how the model works through two case studies. We first describe how we operationalized exploration and exploitation, and our measure of optimal decision performance.

B. OPERATIONALIZATION OF EXPLORATION AND EXPLOITATION VIA TME MONITORING OF SEQUENTIAL SAMPLE VARIANCES

We hypothesize that variability in latency times could be used as a way to operationally define the cognitive states of exploration and exploitation. Specifically, we expect that high variability in latency times is indicative of seeking, responding, and synthesizing information that occurs with exploration, whereas low variability in latency times signifies exploitation.

One method for monitoring latency variability is via a sequential scheme, where the variance of a latency measure is repeatedly estimated from moving windows of data. Specifically, let x_i denote the latency at time i, $i = 2, 3, \ldots, 200$. Then, for some window of data of size $w + 1$, starting at time $i = w + 2$, sequentially calculate

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

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 the sequence of sample variances is less than some threshold h, we declare that the subject has gone from exploration to exploitation.

For this method, one question is how to choose w. There are two considerations: (1) ideally $w + 1$ should be smaller than the smallest length of time that a subject might be in exploration mode when the experiment first starts, and (2) smaller values of w are 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 of $5 \leq w \leq 20$ or so.

A second question is how to choose h. The planned approach will be to subjectively compare how well various values of h differentiate between exploration and exploitation, as determined by various other external measures, such as those from the

EEG, on a training set of data. The value of h that performs best will then be applied to the remaining data.

Finally, there is also a question of whether and how to detect if someone reverts from exploitation back to exploration. One possibility is to continue to monitor the sample variances and, once someone is in exploration mode, should $s_i^2 > h$, conclude that they have reverted back to exploration. However, it may be that we need two thresholds, call them h_1 and h_2, where $h_2 > h_1$, which would work as follows. For someone in exploration mode, they switch to exploitation at time i when $s_i^2 < h_1$, while for someone in exploitation mode, they only switch to exploration at time i when $\boldsymbol{s_i^2 > h_2}$. 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.

C. MEASURE OF REGRET AS A OBJECTIVE MEASURE OF DECISION PERFORMANCE

Regret provides a measure of deviations from the ideal decision path, at any given point in a task. Regret is the difference of a trainee's single trial outcome and the outcome from the ideal decision, given perfect knowledge. Less regret is better; on any given trial, regret can be zero if the trainee selects the best decision. More generally, absolute regret compares the outcome of trainee actions to the outcome generated by playing the optimal policy at each of the n trials. Given K ≥ 2 routes and sequences $r_{,\mathrm{i},1}$, $r_{,\mathrm{I},2}...r_{i,n}$ of unknown outcomes associated with each route i = 1,...K, at each trial, t = 1,...n, trainees select a route I_t and receive the associated outcomes $r_{It,t}$. Let $r_{i,t}^*$ be the best possible outcome possible from route i on trial t (Auer & Ortner, 2010). The regret after n plays $I_1,...I_n$ is defined by

$$R_n = \sum_{t+1}^{n} r_{i,t}^* - \sum_{t+1}^{n} r_{I,t} \ .$$

Regret provides insights in the aggregate over the course of a set of n trials (i.e., total regret) and, when examined, per trial. Regret per trial provides a measure of a trainee's ability to identify the best choice available at a given point in time.

D. USE OF NEUROPHYSIOLOGICAL MEASURES TO PROVIDE INSIGHTS INTO WHY NONOPTIMAL DECISION MAKING OCCURRED

Numerous studies indicate that eye-movement data via eye-tracking technology can provide valuable insights into subjects' attention allocation patterns and underlying cognitive strategies during real-world tasks (Kasarskis, Stehwien, Hickox, Aretz, & Wickens, 2001; Marshall, 2007; Sullivan, Yang, Day, & Kennedy, 2011).

E. CAPTTIM

Figure 1 outlines the main component of CAPPTIM: determining if a trainee's cognitive state is aligned or misaligned with their actual performance. When cognitive state is misaligned with actual performance, it indicates that a training intervention is required. As illustrated in Figure 1, a trainee typically would start in the yellow cell, in which they are in exploration mode and their decision performance is nonoptimal. Ideally, at some point during the task, the trainee transitions to the green cell, in which they are in exploitation mode and their decision performance is optimal, as indicated by low regret. When a trainee's cognitive state is misaligned with actual decision performance, training intervention should occur (orange and red cells). Given that latency variance and regret can be measured in real time, the combination of these two measures can be used as a simple, near-immediate indicator of training intervention.

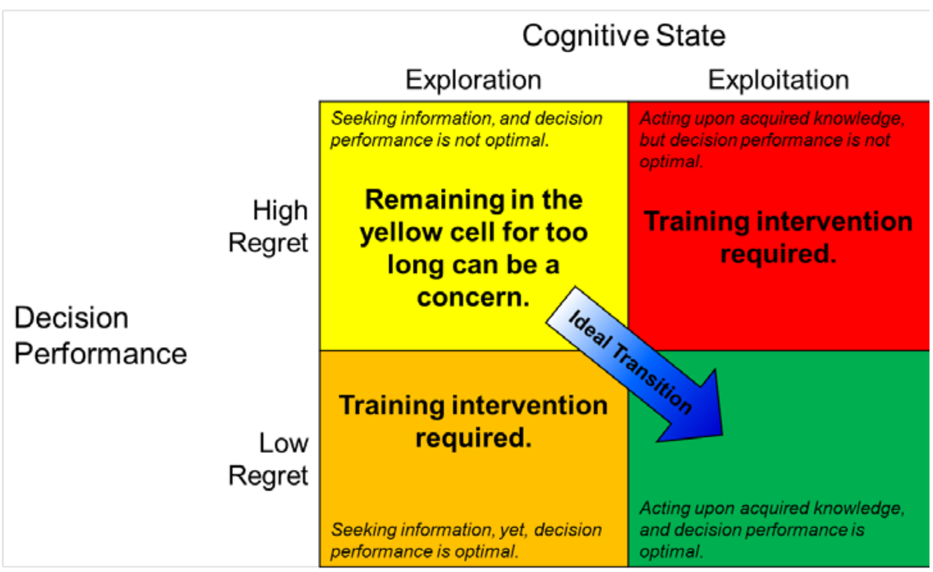


Figure 1. Illustration of the main components of CAPTTIM.

The model determines whether cognitive state, exploration or exploitation, is aligned or misaligned with actual decision performance, as measured by regret. The alignment or misalignment is an indicator of the quality of the decisions and the trainee's mastery of the task. When misalignment occurs, a training intervention is required. Misalignment can occur for several reasons, such as lack of focus on the relevant information, distraction, sleepiness, or high cognitive workload.

Next, the incorporation of neurophysiological measures, such as eye tracking and electroencephalography (EEG), can provide an understanding as to *why* a trainee's cognitive state and actual performance are misaligned (see Figure 2 and Table 1). Understanding why misalignment between cognitive state and decision performance occurred can inform the type of training intervention that should be done. For example, perhaps a trainee is in the red cell simply because they are not attending to the most relevant pieces of information. In this case, an attention allocation intervention could be employed. A trainee in the orange cell may be experiencing an overly high cognitive workload during the task and therefore does not have the cognitive capacity to realize that they are performing well. In this case, an intervention that uses very strong positive feedback could help the trainee realize that they actually have figured out the task. Thus, these initial results suggest that highly efficient and targeted training interventions can occur with the combined use of decision performance, time to make a decision, eye-tracking, and EEG information monitored in real time. In the next section, we illustrate CAPTTIM with two case studies.

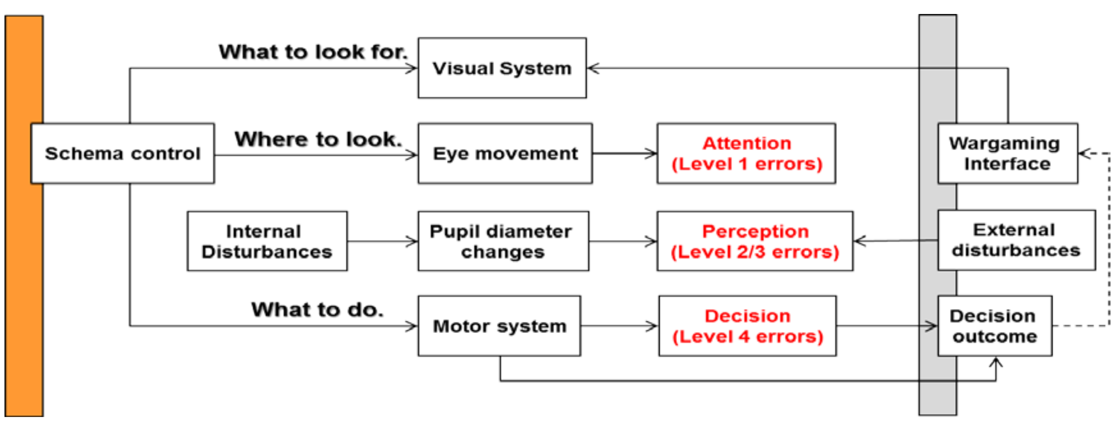


Figure 2. Adapted from Land & Hayhoe (2001), this figure illustrates how neurophysiological data can inform why nonoptimal decision making occurred.

Table 1. Outline of the secondary component of CAPTTIM: targeting the training intervention. Included is a description of each type of nonoptimal, decision-making error and a corresponding possible training intervention.

Error level	**Description**	**Possible Training Intervention**
Attention(Level 1 errors)	Information from eyetracking indicates that the person was not looking at the salient information; therefore, optimal decision making is unlikely to occur.	Attention allocation that directs trainee's gaze to the salient information.
Perception (Level 2 error)	Information from eyetracking indicates that the person glanced at the salient information, but not long enough for it to register in the brain.	Attention allocation that directs trainee's gaze to the salient information.
Perception (Level 3 error)	Information from eyetracking indicates that the person looked at the salient information, and long enough for that information to register in the brain. However, EEG data shows that the person is experiencing one or a combination of the following: high cognitive workload, frequent distraction, or sleepiness.	Different training interventions depending on the EEG data. High cognitive workload: restart the task at a lower level of difficulty. Distraction: Focus the trainee's attention on the task; reduce distraction in the surrounding area. Sleepiness: Trainee should resume the task at a later time.
Decision (Level 4 error)	This error occurs due to the person incorrectly using past experience or preconceived notions in making their decisions. Information from eyetracking and EEG rule out level 1-3 errors. The person is looking at the salient information and they are not experiencing high cognitive workload, distraction, or sleepiness.	Increasingly stronger visual/audio cues to the trainee that their current strategy is not optimal. Strong, immediate, positive feedback when the trainee makes optimal decisions.

F. ILLUSTRATION OF CAPTTIM WITH CASE STUDIES FROM THE CONVOY TASK

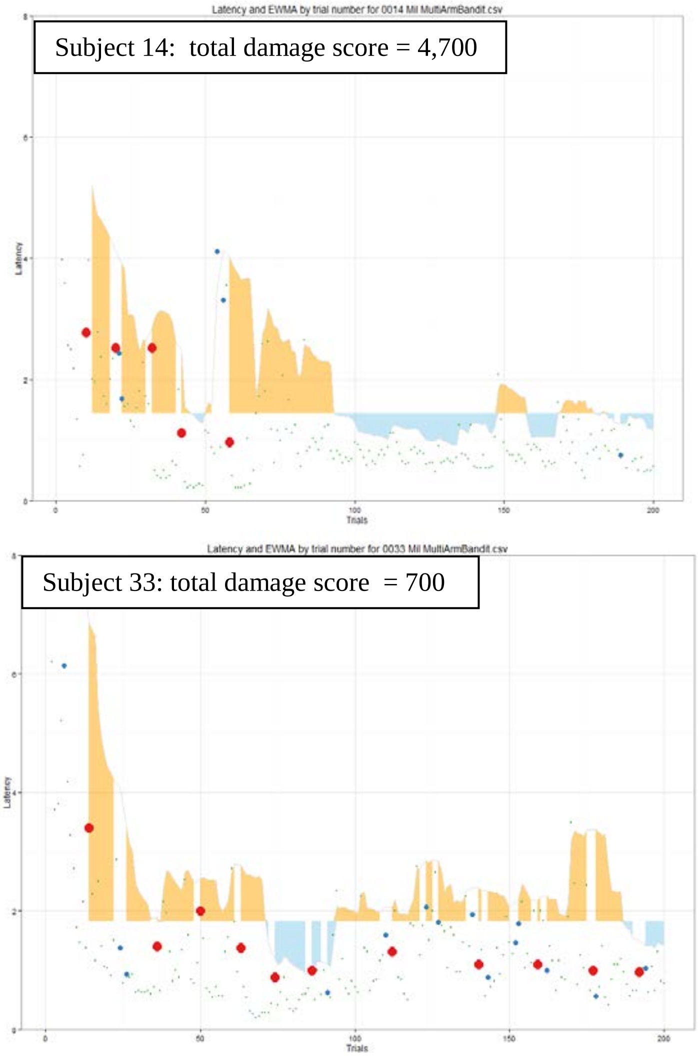
In Kennedy, Nesbitt, and Alt (2014), we developed and tested a simple wargame called the convoy task on 34 subjects, all of whom were military officers. In the convoy task, subjects see four identical roads and are instructed to select the route on which to send their convoy (see Figure 3). Their goal is to have the highest total damage score by maximizing the damage to enemy forces, while minimizing the friendly damage accrued over all trials. Through trial and error, subjects learn which routes have the best long-term payoffs in damage. On each trial, the subject is provided immediate feedback in the form of three separate pieces of information: a reward, a penalty, and a running total. The reward—the number of enemy forces damaged—is called Enemy Damage. On any given trial, enemy damage ranges from 50 to 100 damage. The penalty—the number of friendly forces damaged—is called Friendly Damage. Depending on the route chosen, friendly damage ranges from 0 to –1,250 damage. The running total is called Total Damage, defined as the previous trial's value of Total 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. Subjects begin the task with 2,000 damage. The main outcome variable is Total Damage at the end of the 200 trials. A subject selects routes until the end, not knowing that the task will complete after 200 trials. 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 incorporate information indexed by trial into their estimate.

Each route has its own scripted, ordered set of specified values. For example, every subject will find that the third time they pick route *1*, it returns +100 enemy damage and –150 friendly damage. Even though these returns by route are set and are the same for each trainee, the games will progress differently due to the divergence of route selection between subjects.

Figure 3. Screen shot of the convoy task in piloting; a typical subject's view of the task. We see that the trainee's last choice caused 100 damage to the enemy (Damage 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 2,750. A positive Accumulated Damage value is desirable to the trainee. Notice that four routes are represented by the same image.

G. SEQUENTIAL DETECTION METHOD: USING LATENCY DATA TO DETERMINE EXPLORATION VS. EXPLOITATION COGNITIVE STATES

As illustrated in Figures 4a and 4b, we successfully used variability in trial-by-trial latency time to detect periods of exploration and exploitation cognitive states. A single explore/exploit latent threshold was developed for each subject, derived from twice the standard deviation above and below all latency times for 0 or 50 friendly damage (i.e., the baseline latency time) for that subject. Therefore, exploration was defined as trials in which the latency time was at least two standard deviations (SD) higher than the baseline latency time. Exploitation was defined as two SD lower than the baseline latency time. Note that these definitions do not take into account actual decision performance, but solely the subject's cognitive state at a given time in the task. Figures 4a and 4b depict two distinct patterns of exploration and exploitation. Figure 4a depicts an optimal exploration to exploitation transition (subject 14), whereas Figure 4b illustrates a pattern of primarily exploration throughout most of the task (subject 33).

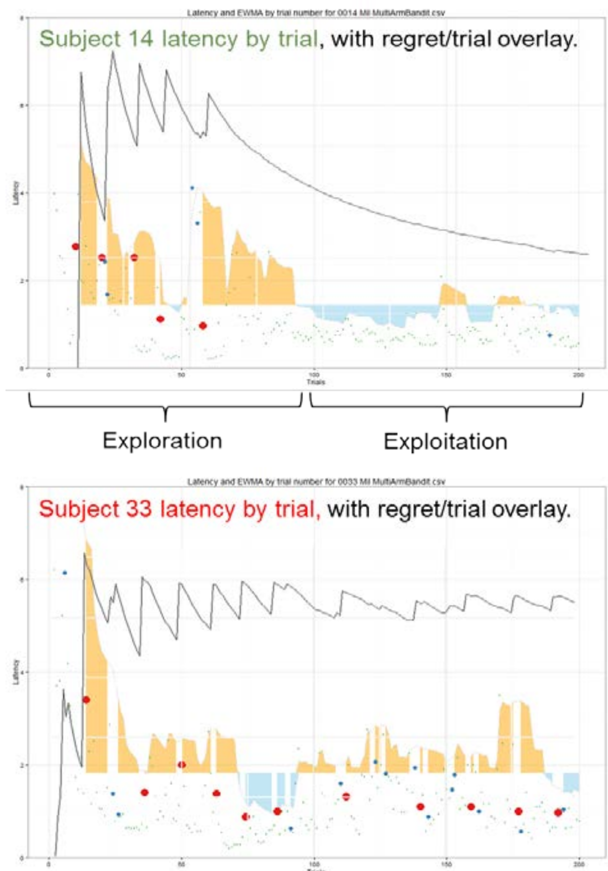


Figures 4a and 4b. Use of sequential sample variances in latency times to determine exploration and exploitation cognitive states. Shaded orange regions indicate periods of exploitation; shaded blue regions indicate periods of exploitation.

H. COMBINING SEQUENTIAL DETECTION METHODS WITH REGRET

The combination of trial-by-trial information regarding the subject’s current cognitive state (exploration or exploitation) with actual performance (measures of regret) provides insights into whose cognitive state is aligned with actual performance. Across the 34 subjects who completed the convoy task, clear patterns of cognitive alignment and misalignment are seen. We illustrate two of these patterns, exhibited by subjects 14 and

33, in Figures 5a and 5b. In Figures 5a through 5d, we see that although subjects 14 and 33 show distinct differences in cognitive state, their cognitive state is aligned with their measure of regret. Subject 14 goes through a period of exploration until about trial 90, at which point they are predominantly in exploitation mode. Consistent with this cognitive state pattern, subject 14's regret is quite high until about trial 90, at which point it begins to steeply decrease. Recall that lower regret means that the subject's decisions are verging towards the best possible decision. Thus, when subject 14's cognitive state is in exploration mode, their regret is correspondingly high. When their cognitive state transitions to exploitation, their regret consistently decreases. In contrast, subject 33 maintains an exploration cognitive state throughout most of the task and, correspondingly, their regret is consistently high throughout the task.



Figures 5a and 5b. Figures 5a and 5b illustrate the concordant pattern between subject's cognitive state and their actual decision performance, as measured by regret, for two different subjects. Regret across the 200 trials is denoted by the black line.

We then examined subject 33's eye gaze and EEG data for indicators as to why subject 33 showed a nonoptimal pattern and poor decision performance. As outlined in Table 2, eyetracking data indicates that subject 33 had a similar eye gaze pattern as the overall sample and that this subject was correctly focusing on friendly damage to a much greater extent than total damage or enemy damage.

Table 2. Comparison of subject 33's eye gaze pattern compared to the overall sample.

	Total Damage	**Friendly Damage**	**Enemy Damage**	**Routes**
Mean gaze time (SD), (%)	5.49 (12.47)	16.73 (14.87)	6.55 (6.40)	71.23 (19.86)
Subject 33	2.90	13.96	7.78	75.26

Figure 6 illustrates the utility of combining neurophysiological and behavioral measures. Subject 33's EEG data indicates that there were several periods throughout the task when they experienced high cognitive workload. Note that the peaks in latency time in the first several trials, and between approximately trials 160 to 170, overlap and/or precede peaks in periods of high cognitive workload. This subject, however, was also frequently distracted and was minimally engaged in the task. Given insight into the subject's cognitive state throughout the task, it is not that surprising that subject 33 remained in an exploration state, had high regret, and scored 700 in total damage, which was well below the average of 2,402.94.

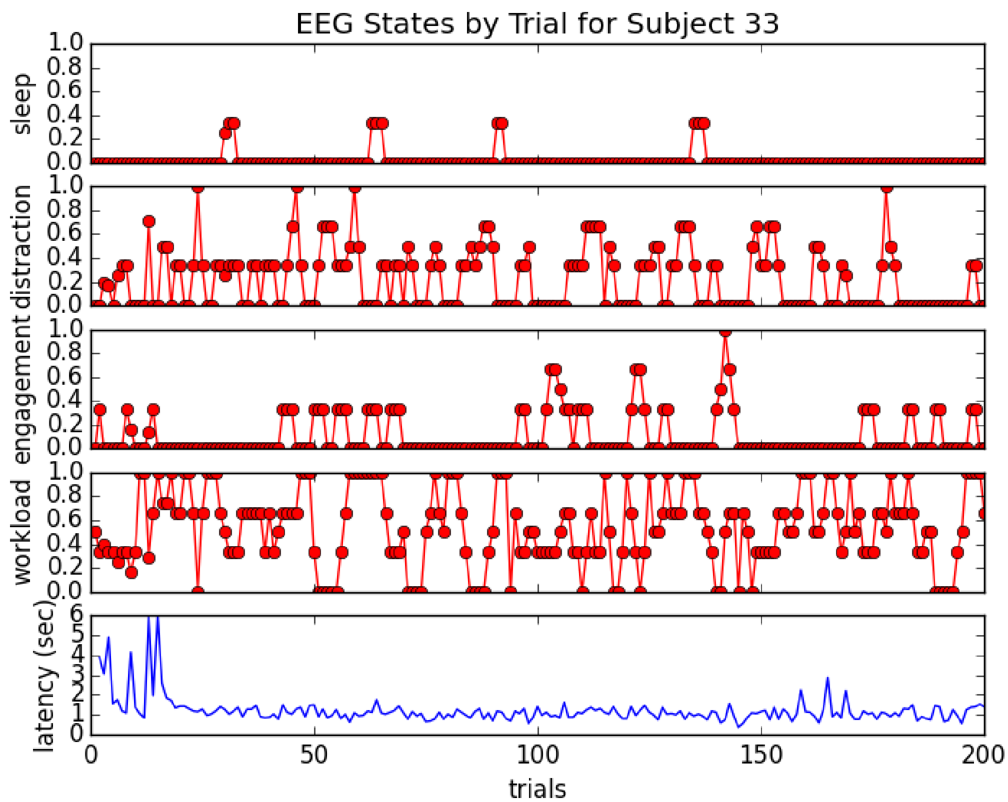


Figure 6. The proportion of time that subject 33 experienced sleepiness, distraction, high engagement, or high cognitive workload on a given trial. Latency per trial is depicted as the blue line.

II. SUMMARY

The purpose of this paper was to use case studies to illustrate CAPTTIM and its potential impact on current military training. CAPTTIM uses quantitative statistical methods and objective neurophysiological measures to complete the following actions in real time: (1) characterize a trainee's cognitive state as either exploration or exploitation, (2) determine whether cognitive state is aligned or misaligned with actual performance, and (3) indicate ways in which the training intervention can be targeted to address why cognitive misalignment occurred. Because latency times and decision performance measures, such as regret, are simple behavioral measures that easily can be programmed into training software, this process can be completed in real time, with near-immediate notification that a training intervention is required. Neurophysiological measures, such as eyetracking and EEG, also are measured continuously and in real time, suggesting the potential for a near-immediate, targeted training intervention. Because of these characteristics, CAPTTIM has the potential to improve current military training efficiency and effectiveness.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

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

Kasarskis, P., Stehwien, J., Hickox, J., Aretz, A., & Wickens, C. (2001). Comparison of expert and novice scan behaviors during VFR flight. In *Proceedings of the 11th International Symposium on Aviation Psychology*.

Kennedy, Q., Nesbitt, P., & Alt, J. (2014). Assessment of cognitive components of decision making with military versions of the IGT and WCST. *Human Factors and Ergonomics Society 2014 International Annual Meeting*, October 27-31, Chicago, IL.

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, *41*(25), 3559–3565.

Lopez, T. (2011). Odierno outlines priorities as Army Chief. *Army News Service*. Retrieved from http://www.defense.gov/News/NewsArticle.aspx?ID=65292s

Marshall, S. P. (2007). Identifying cognitive state from eye metrics. *Aviation, Space, and Environmental Medicine*, *78*(Supplement 1), B165–B175.

Sullivan, J., Yang, J. H., Day, M., & Kennedy, Q. (2011). Training simulation for helicopter navigation by characterizing visual scan patterns. *Aviation, Space, and Environmental Medicine*, *82*(9), 871–878.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Research Sponsored Programs Office, Code 41
   Naval Postgraduate School
   Monterey, California

4. Richard Mastowski (Technical Editor)..1
   Graduate School of Operational and Information Sciences (GSOIS)
   Naval Postgraduate School
   Monterey, California

5. Dr. Virginia Pasour..1
   U.S. Army Research Office (ARO)
   Mathematical Sciences Division
   4300 S. Miami Blvd.
   Durham, NC 27703

6. Senior Lecturer Quinn Kennedy..1
   Operations Research Department
   Naval Postgraduate School
   Monterey, CA 93943

7. MAJ Peter Nesbitt, USA..1
   Operations Research Department
   Naval Postgraduate School
   Monterey, CA 93943

8. LTC Jon Alt, USA 1
   Operations Research Department
   Naval Postgraduate School
   Monterey, CA 93943

9. Professor Ronald D. Fricker, Jr...1
   Operations Research Department
   Naval Postgraduate School
   Monterey, CA 93943

Appendix D
References

[1] U.S., Department of the Army Training and Doctrine Command. *TRADOC Pamphlet 525-3-1, The U.S. Army Operating Concept: Win In a Complex World*. Washington DC: Government Printing Office, October 2014.

[2] Antoine Bechara et al. "Insensitivity to future consequences following damage to human prefrontal cortex". In: *Cognition* 50.1 (1994), pp. 7–15.

[3] 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." In: *Journal of experimental psychology* 38.4 (1948), p. 404.

[4] Quinn Kennedy, Peter Nesbitt, and Jon Alt. "Assessment of Cognitive Components of Decision Making with Military Versions of the IGT and WCST". In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 58. 1. SAGE Publications. 2014, pp. 300–304.

This page intentionally left blank.

Appendix E
Glossary

ARO	Army Research Office
CAPTTIM	Cognitive Alighnment With Performance Targeted Training Intervention Model
IGT	Iowa Gambling Task
TRAC	Training and Doctrine Command Analysis Center
WCST	Wisconsin Card Sorting Test