2018

An Agent Based Model to Assess Crew Temporal Variability During U.S. Navy Shipboard Operations

Kevin Muhs
University of Central Florida

AN AGENT BASED MODEL TO ASSESS CREW TEMPORAL VARIABILITY
DURING U.S. NAVY SHIPBOARD OPERATIONS

by

KEVIN S. MUHS
B.S. University of Washington, 1987
M.S. Naval Postgraduate School, 1995
M.S. Naval Postgraduate School, 1995

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Industrial Engineering and Management Systems
in the College of Engineering and Computer Sciences
at the University of Central Florida
Orlando, Florida

Spring Term
2018

Major Professor: Waldemar Karwowski

ABSTRACT

Understanding the factors that affect human performance variability as well as their temporal impacts is an essential element in fully integrating and designing complex, adaptive environments. This understanding is particularly necessary for high stakes, time-critical routines such as those performed during nuclear reactor, air traffic control, and military operations. Over the last three decades significant efforts have emerged to demonstrate and apply a host of techniques to include Discrete Event Simulation, Bayesian Belief Networks, Neural Networks, and a multitude of existing software applications to provide relevant assessments of human task performance and temporal variability. The objective of this research was to design and develop a novel Agent Based Modeling and Simulation (ABMS) methodology to generate a timeline of work and assess impacts of crew temporal variability during U.S. Navy Small Boat Defense operations in littoral waters.

The developed ABMS methodology included human performance models for six crew members (agents) as well as a threat craft, and incorporated varying levels of crew capability and task support. AnyLogic ABMS software was used to simultaneously provide detailed measures of individual sailor performance and of system-level emergent behavior. This methodology and these models were adapted and built to assure extensibility across a broad range of U.S. Navy shipboard operations.

Application of the developed ABMS methodology effectively demonstrated a way to visualize and quantify impacts/uncertainties of human temporal variability on both workload and crew effectiveness during U.S. Navy shipboard operations.

This dissertation is dedicated to the entirety of my dearly loved family in appreciation of the steadfast support, patience, and love provided over the course of our lives together and particularly during my time as a student. Tammy, Anna, Bradley, Stephany, and Edith have been the bedrock of inspiration and encouragement needed to complete this journey. Life is good!

ACKNOWLEDGMENTS

I would like to acknowledge my deep appreciation of several individuals for their contributions to this dissertation and the support provided in my pursuit of this degree.

1. The chair of the dissertation committee and my dissertation adviser, Dr. Waldemar Karwowski, whose tremendous insight, guidance, and imparted scholarly discipline ensured that this dissertation met its potential.
2. The remaining members of my committee, Dr. Peter Hancock, Dr. Glenn Sjoden, and Dr. Ahmad Elshennawy provided invaluable guidance that greatly enriched the depth and fidelity of this dissertation.
3. Ms. Sabrina Miller, a tremendous research librarian, whose knowledge of defense generated research and patience in answering my many requests were sincerely appreciated.
4. Mrs. Anne Muhs, a lifelong literature teacher, whose patience and artful persistence in reviewing and re-reviewing the many edited versions of this manuscript have resulted in an effective and proper communication of the work completed. Her proactive willingness in the conduct of these reviews is greatly appreciated and very meaningful to me personally.

In addition, I would also like to thank the Office of Naval Research for partially funding this research through the following two grants:

1. Grant N00014-14-1-0777
2. Grant N00014-13-C-0066

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE: INTRODUCTION

1.1 Background

United States Navy sailors are consistently relied upon to provide a high level of performance in extreme conditions and over extended periods of time. This reliance requires Navy leadership, design planners, and operational commanders to take many human factors into account across a broad range of phases, spanning from shipbuilding design to shipboard training and operations. Weaknesses in planning, preparing, and executing in any of these phases can result in catastrophic loss of human lives, damage to major systems, and/or significant financial costs. In 2017, the potential for these negative outcomes was clearly demonstrated as the U.S. Navy suffered three collisions and one grounding in the Western Pacific. In each of these incidents, the time to complete required preventative and/or mitigating actions played a critical role in contributing to both the occurrence and impact of the event (Davidson, 2017). Two of the collisions directly resulted in major shipboard flooding, loss of critical systems, and the death of seventeen U.S. Sailors. As a result, renewed interest by U.S. Navy leadership has emerged that emphasizes effective development of human factor predictive performance standards through the collection and modeling of human variability. A keen understanding of one of these human performance factors, temporal variability in the completion of tasks, is necessary to assure attainment of the stated U.S. Navy operational performance initiatives, maintain effective readiness of the force, and efficiently manage crew training.

Agent-based modeling and simulation (ABMS) may provide an effective method of evaluating the effects of this human temporal variability and resulting shipboard operational impacts. For example, human performance models for Officers and Enlisted, interacting with one another as well as with shipboard communication, navigation, surveillance, and fire control technologies, may prove beneficial in assessing both the individual behavior of these sailors and the overall emergent behavior of shipboard operations.

1.2 Goals and Objectives

The goal of this research is to develop an ABMS method to investigate the effects of human performance variability on U.S Navy sailor actions for a unique, time-critical component of shipboard operations. Once developed, this ABMS method could be expanded to evaluate both battle station and routine sailor activities to support operations. In order to exercise the capabilities of the proposed ABMS approach, a test scenario was developed to investigate the impact of sailor temporal variability on the crew's ability to defend the ship against a small boat attack while operating in littoral waters. This example provided an opportunity to 1) explore a problem with an expected impact at the ship system-wide level; 2) address a relevant U.S. Navy safety issue; and 3) leverage the U.S. Navy Human System Integration (HSI) task database, a comprehensive listing of crew-performed tasks consisting of 78 attributes for each task.

Specific application of the ABMS method was then completed by using subject matter expert (SME) input and HSI database mining to identify representative times for each of the tasks within the small boat defense scenario of interest. Once representative times were identified, the ABMS approach was used to assess the watch team temporal variability for each task within a given scenario. Cumulative task times were then summed to identify overall scenario completion times for comparison with a projected critical time for completion. This approach supported the following research questions (RQ):

- RQ1. How can ABMS be used to predict scenario temporal outcomes through workflow evaluation and assessment of U.S. Navy shipboard personnel capabilities?
- RQ2. Given a critical/desired response time and parameterized crew temporal variability, how well can ABMS be used to determine the likelihood of exceeding the designated time?
- RQ3. How can the ABMS approach be used to determine the impact of performance influencing factors, such as sailor capability and task support, on crew temporal variability in the performance of a defined scenario?

CHAPTER TWO: LITERATURE REVIEW

A simulation model is built upon beliefs and assumptions about the behavior of an actual system (Garrido, 2001). Thus, agent based modeling and simulation allows the opportunity to gain insight into the working relationships and behaviors of sailors in the performance of their duties onboard ship, and helps to assess the factors affecting sailor temporal variability impacts. This chapter presents an overview of U.S. Navy small boat defense operations in littoral waters and the ABMS design approach used in developing representative crew temporal variability outcomes and impacts. In addition, previous work on the factors and impacts of human temporal variability completed by the author and published in Muhs, Karwowski, and Kern (2018) is presented.

2.1 U.S. Navy Littoral Water Operations

United States Naval forces routinely conduct transits in waters lying along the shores of foreign nations. These areas, known as littoral waters, present a host of unique force protection challenges. The mechanisms employed by the U.S. Navy to address such challenges, as well as to successfully maintain routine operations of the ship, begin with the "on duty" watch team. This team consists of rotationally assigned, qualified specialists who operate the ship continuously and assure that the routine functions of the ship run smoothly. A secondary function of the watch team is to respond to emergencies and force protection issues arising on the ship or involving other ships. On a typical U.S. Navy vessel, these personnel keep watch on the bridge and over the running machinery throughout the ship. The bridge is staffed 24 hours a day and typically consists of six to ten members responsible for safe navigation and operation of the vessel. This watch team is led by a watch officer, who reports to the Commanding Officer. Below the bridge is the combat information center (CIC), manned by a watch team of six to ten officers and enlisted specialists responsible for the weapons system. The CIC includes a radar operator who monitors ships within range and a fire control technician who monitors displays and assigns electronic tracking tags to each of the contacts identified by the radar operators, lookouts, or other members of the watch team. Lookouts are typically stationed at the back and near the front of the ship, on or close to the bridge. Lookouts and other portions of the bridge watch team constantly scan the horizon with binoculars to back up the radar operator in case of missed small boats approaching the vessel. They may also identify inadvertently tracked waves caused by a heavy sea, which can present as contacts. The entire watch team is capable of being in constant contact over various types of radio communications. While in littoral waters or the open sea, bridge-to-bridge radios using a common frequency are used to communicate in real time with other vessels.

Understanding and modeling these complex watch team interactions and the temporal variability parameters defining them requires a dynamical systems approach methodology along with the application of complex and adaptive modeling techniques.

2.2 Agent Based Modeling and Simulation

An agent based model (ABM) contains one or more autonomous agents that can perceive their environment, exchange information, make operational decisions, and act based on those decisions (North & Macal, 2007). These mechanisms of response are representative of the real world processes and interactions that exist between crew members in the performance of their duties onboard every U.S. Navy ship.

2.2.1 Introduction

Agent Based Modeling is a methodology for mapping the actions and interactions of autonomous individuals into a computer program with a view to assessing their effects on the system as a whole (Zheng, et al., 2013). In short, ABM expresses real world processes in terms of algorithms and mathematical formulas that are implemented as a code in a programming language (Baqueiro, Wang, McBurney, & Coenen, 2009).

Baqueiro et al. (2009) also identify that ABMS allows experimental designs that test the developed models and theoretical frameworks under different scenarios with different parameter configurations. These experiments provide designers with an insight to certain aspects of a complex system that would not otherwise be possible using mathematical analysis alone. Designers in their application and use of ABM are typically seeking to accomplish prediction (making prognoses); verification (to determine if designed models are correct); validation; training (improving skills); and/or increased knowledge of subjects or domain (Wooldridge, 2009). The typical agent based models consist of three elements (Macal & North, 2014):

- Agents: Their attributes and behaviors.
- Agent relationships and methods of interaction.
- Agents’ interaction with external environment/influences and other agents.

In the design of the ABM, these elements have dynamic and coherent relationships within the sphere of their influence and the environment in which they exist.

2.2.2 Agents and ABM Design

An agent is the basic component of any ABM and represents an autonomous knowledge based system that perceives the environment, is capable of reasoning about a given situation, makes decisions independently, and executes tasks to accomplish the goals of a mission (Mandal, Han, Pattipati, & Kleinman, 2010). However, this description of an agent does not translate within the ABM research community to a universally accepted definition for the term "agent" (Macal & North, 2014). Different modelers look at agents from different perspectives. For example, Bonabeau (2002) defines an agent as an independent component whose behavior can vary from primitive, reactive decision rules to complex, adaptive intelligence. In contrast, Mellouli et al. (2004) define an agent to be any independent component with adaptive behavior and an ability to learn from its environment and change the behavior in response. Jennings (2000), on the other hand, used a computer science based view of an agent to emphasize autonomous behavior.

Independent of the definition used, an agent's considered characteristics may be different as well, depending on the real world system being modeled (Wooldridge, 2009). In their work, Macal and North (2014) identify well-established agent characteristics to support practical modeling based on how agent models are built and described. The characteristics are as follows:

1. Autonomy: An agent is autonomous, self-directed, and independently functioning in its environment and interactions with other agents.
2. Modularity: Agents are modular or self-contained. They are identifiable, discrete entities with a set of attributes, behaviors, and decision-making capability.

3. Sociality: An agent is social, interacting with other agents.
4. Conditionality: An agent has a state that varies over time.

Agents may also have additional properties, which may or may not be considered requisite properties for a given modeled system. In addition to characteristics, the type of agent may also be specified. Agents working in a distributed multi-agent environment can be categorized as either benevolent or self-interested. Benevolent agents work together toward a common goal, whereas self-interested agents work independently to achieve their own goals. Additionally, in models that require an understanding of the agent's decision-making process, the agents can be categorized as one of four types (Meirina, Levchuk, & Pattipati, 2003):

1. Logic Based Agents: Logical deduction determines the agent's decisions.
2. Reactive Agents: Agent decision making is implemented in a direct mapping format from stimulus to action.
3. Belief-Desire-Intention Agents: Agent decision making depends on manipulation of data structures that represent beliefs, desires, and intentions of the agent.
4. Layered Agents: Agent decision making is based on layered software programs that represent explicit reasoning about the environment.

Once the characteristics and type of agent(s) are identified in the modeled system, they are dynamically coupled with their environments as well as with other agents in the design and development of the ABM.

2.2.3 Design and Development of Agent Based Models

The general design of a well-developed ABM interweaves three interdependent stages, which are described by Fishwick (1997) as follows:

1. Designing the model: Constructing a model that is representative of the real system under investigation is the goal of this stage. This is accomplished by using information and data collected from real world observations, in the form of numerical values and abstract concepts, to build a mathematical model.
2. Executing the model: In this stage, the mathematical is converted into computer algorithms, which are then executed to produce data in the form of numerical values.
3. Analyzing the outcome: This stage compares simulated data with data produced by the mathematical model.

Fishwick (1997) states that these three stages work closely with one another and that the entire agent based modeling and simulation process comprises a finite number of iterations. Data generated by executing these models can then be compared to data observed from other independent sources to support validation. If the ABM developed data set does not conform to the real world observations, altered assumptions can be used to repeat the process until a valid model is obtained.

2.3 Temporal Variability in Human Performance

In human performance, temporal variability is ubiquitous. The effects of this variability can be seen in the perceptual, cognitive, and physical dimensions of human performance when interacting within complex socio-technical systems (STS). Understanding the factors that affect performance variability as well as their temporal impacts is an essential element in fully integrating and designing humans into complex environments. As a result, accurate prediction of the factors affecting temporal variability within the context of individual task performance, as well as the development, refinement, and use of reliable tools in assessing this variability, has been a major focus in research for well over fifty years. Over this time, significant understanding of the individual elements and organizational factors that impact human temporal variability has been gained through discerning research and broad coverage in literature. Components of the research (Maynard, Stegemerten, & Schwab, 1948; Hick, 1952; Hyman, 1953; Fitts, 1954) have been generalized and extremely far reaching in the field of Human Factors and Ergonomic (HF/E) Sciences, whereas others (Chan, Shum, Law, & Hui, 2003; Chen & Joyner, 2006; Stanton & Baber, 2008) have been exceedingly limited in their scope and application. This variance is not unexpected given the broad desire to create both generalized and adaptive rules to human response variability as well as a recognition that context specificity of the task plays a significant role. The variance, as seen in the literature, is also indicative of the shift in human sciences from prescriptive to descriptive models in terms of a rational performance standard in modeling the "actual behavior" as described by Rasmussen (Rasmussen, 1997). Over time and based on the diverse uses of human response data, an ambiguous and myopic divide has appeared in the literature between cognitive and physical human models. However, it is widely accepted that in order to fully appreciate the aggregate temporal variability in human task performance, full and integrated consideration must be given to both the cognitive and physical components of human performance as well as to any interplay that exists between them. Fortunately, recognition of, and advocacy for the need of combined cognitive-physical models is seen to be gaining momentum within research community literature (Zhang, 2003; Badler, et al., 2005; Fuller, Reed, & Liu, 2010; Marras & Hancock, 2014). In addition, task analysis tools have begun limited integration of both the cognitive and physical response aspects of human performance (Allender, 2000; Gore, Hooey, Foyle, & Scott-Nash, 2008; Wong, Walters, & Fairey, 2010). This integration synthesizes the nature and implications of biological factors, learning ability, and organizational task design as well as the respective impacts on human temporal variability from either an inter- or intra-individual basis. As a consequence, meaningful consideration must be given to the wide breadth of research on the mechanisms, taxonomy, and time responses of human task performance as well as to the factors that influence the response itself, if one is to fully understand the growing body of knowledge on this topic. Thus, a systematic review of literature was completed to develop an understanding of the state of research on human response temporal variability in the performance of tasks, their cumulative impacts, and the factors that affect them.

2.3.1 Temporal Variability Literature Review Method

This systematic literature review was carried out according to the guidelines of Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Moher, Liberati, Tetzlaff, & Altman, 2009). No pre-established or registered protocol existed for this review. The review protocol developed for this systematic review was designed to reduce the possibility that the review would be influenced by research expectations. Protocol development specified the development of research questions and a search strategy.

2.3.1.1 Question Development

Based on the objectives of this systematic review, the following questions were derived and form the basis for this literature review:

1. How has the current research of human temporal response evolved?
2. How can current research of human temporal response be classified?
3. What is the current state of human temporal response research with respect to the identified classification architecture?
4. What can be learned from current human temporal response research that will lead to topics for further investigation?

2.3.1.2 Search Strategy/Execution

A formal search strategy for the review was used to find a comprehensive population of scientific papers relevant to answer the identified research questions. The formal definition of this search strategy allowed the formation of a replicable and open review of external literature. The search strategy consisted of defining the search space and vetting process to be used in identifying relevant material. Current and seminal literature in the field of human temporal response including journal articles, textbooks, proceedings, grey literature, and conference presentations were considered key spaces for this review.

During the search phase, well-known and heavily cited articles were used to develop an initial key word search list resulting in over 35 key word combinations as shown in Fig 1. These key word combinations were then used with popular database search tools to include EBSCO Host, Compendex, IEEE XPlore, Web of Science, DTIC online, PsycInfo, Google Scholar, and ProQuest Dissertations & Theses. A series of key word search term variations, derived from the originally identified works and relevant search articles, were then conducted using the same database tools. This method resulted in a narrowing of the focus to identify the components impacting human temporal variability in task performance.

Table 1: Search Term Key Word Combination

Analytic Network Process	Human perception time response	Human task performance
Cognitive ability	Human performance	Methods-Time-Measurement
Cognitive response	Human performance assessment	Performance influencing factors
Cognitive simulation	Human performance distributions	Performance shaping factors
Human ability	Human physical response time variation	Predetermined time response
Human attention time response	Human reliability	Probabilistic risk assessment
Human causality	Human Reliability Analysis	Psychomotor time response variation
Human computer interaction	Human response	Skill-knowledge-rule model
Human factors	Human response model	Socio-technical systems
Human failure	Human response temporal factors	Taxonomy of human abilities
Human fatigue impacts	Human response time distribution	Work measurement and time standards
Human information processing	Human task loading	Workload prediction

In addition, to assure adequate insight into Department of Defense (DoD) specific research, a governmental research librarian provided technical assistance and key word guidance in identifying representative Defense Technical Information Center (DTIC) material. This search methodology resulted in the identification of over 1700 unique works that contained topical content. After retrieving the articles and isolating an unduplicated population, relevant scientific papers were then selected using a formal screening process incorporating predetermined inclusion and exclusion criteria. Inclusion criteria required the research to (a) be written in

English; (b) be peer reviewed, cited grey literature, and/or Department of Defense conducted research available via electronic databases; (c) depict graphs, charts, equations, and/or tables delineating human temporal response variability for a specified taxonomic structure developed for this review; (d) identify, describe, or use empirical and/or modeled methods to quantify and/or compare variability in human temporal response. Criterion (c) means that initially identified research that focused on only qualitative assessment of human performance influencing factors and/or simply compared temporal variability assessment methods were excluded. Other exclusion criteria were: (a) papers written in any language other than English; (b) book chapters; (c) papers which upon review were not related to the research questions; (d) opinions, viewpoints, anecdotes, letters, and editorials. The study selection process and number of studies selected at various stages is summarized in Figure 1, and have been identified in chronological order of publication in Figure 2.

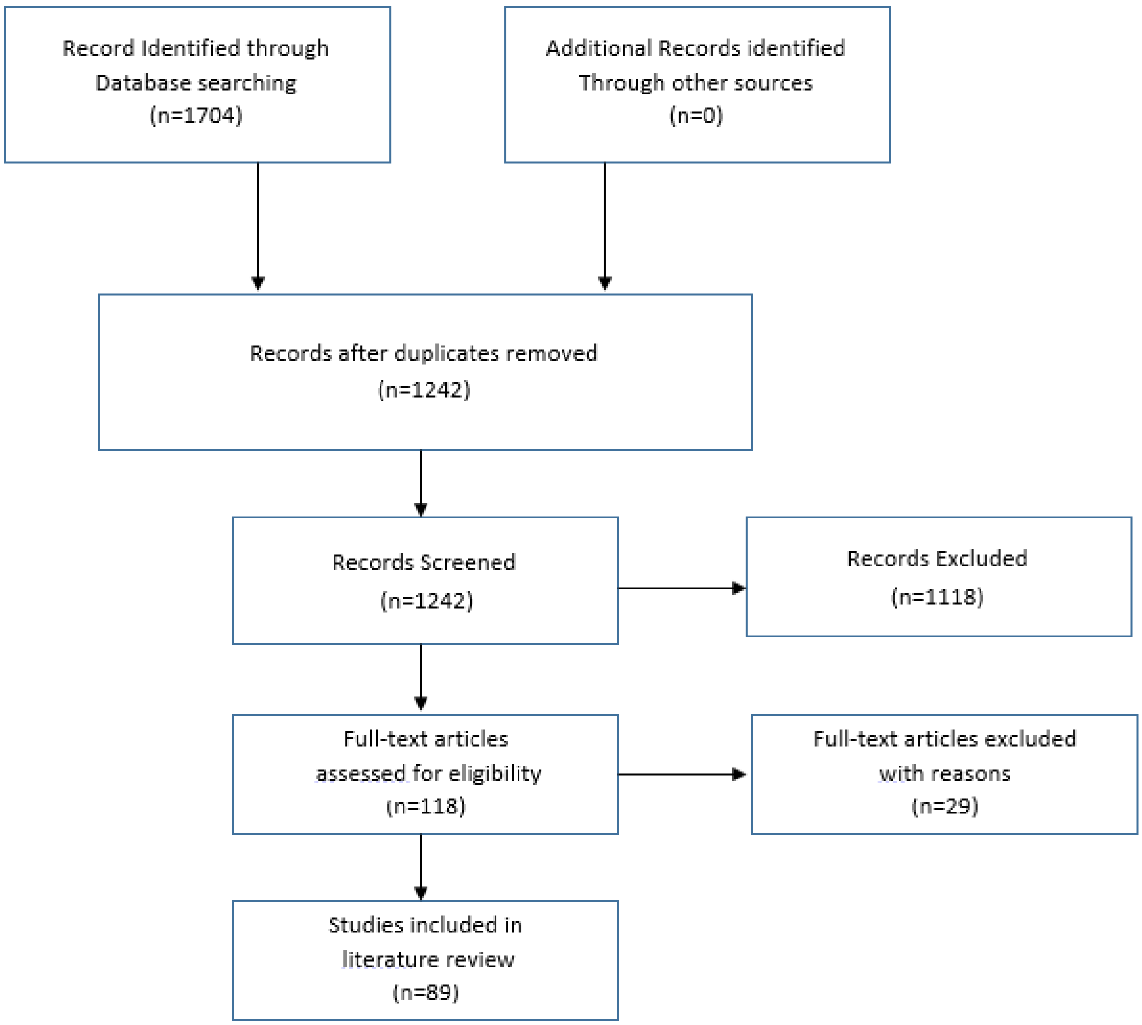


Figure 1: Literature Review Article Selection Process

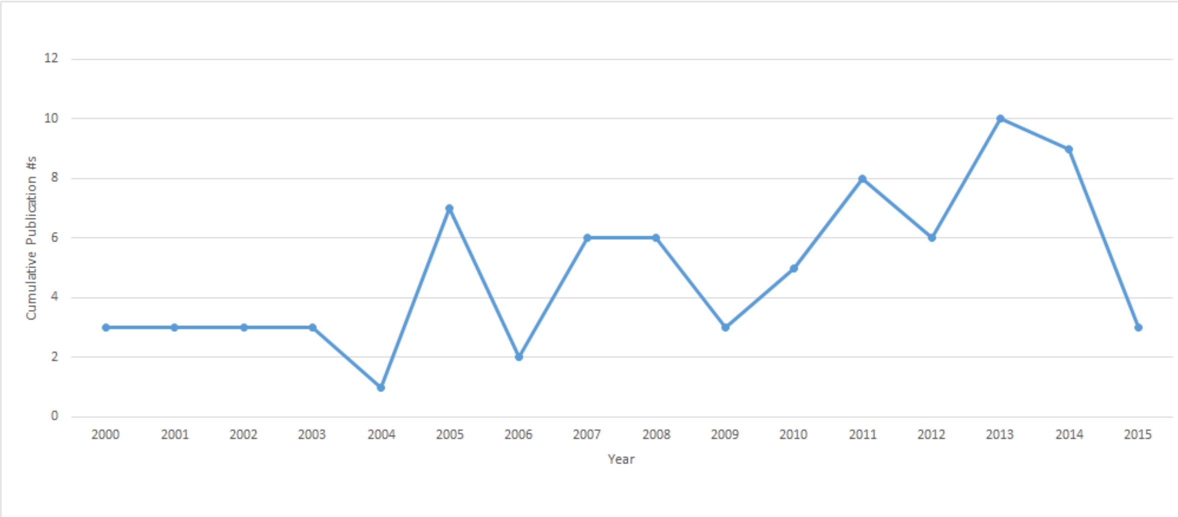


Figure 2: Number of Publications by Year (16 Articles Prior to 2000)

Although not formally established for the study, in practice the year 1998 was typically used as the lower bound for determining the currency of research for the study. However, multiple searches were conducted without regard to timeframe to help identify heavily cited and seminal research in the field of human temporal response. In addition, the ability to work closely with a governmental research librarian resulted in identifying multiple relevant Department of Defense research studies over a span of the last 50 years that influenced the evolution and consideration of human temporal variability within the military. Systematic, all-inclusive searches were continued through the middle of 2015.

2.3.1.3 Data Synthesis and Analysis

Categorization of current research, heavily cited articles, and seminal works to provide historical context was completed by parsing and arranging by commonalities. Selected articles were binned and ordered into five taxonomies (human information processing, psychomotor, physical, performance influencing, and modeling) and arranged by date as well as relevance and content. Topical context was then used to synthesize information in a manner that built upon itself in providing the reader a complete picture of the framework and nature of components contributing to temporal variability in human task performance. This systematic review strategy and process retrieved a combination of 89 relevant and unique scientific peer reviewed papers and military technical reports which have been identified by category in Figure 3.

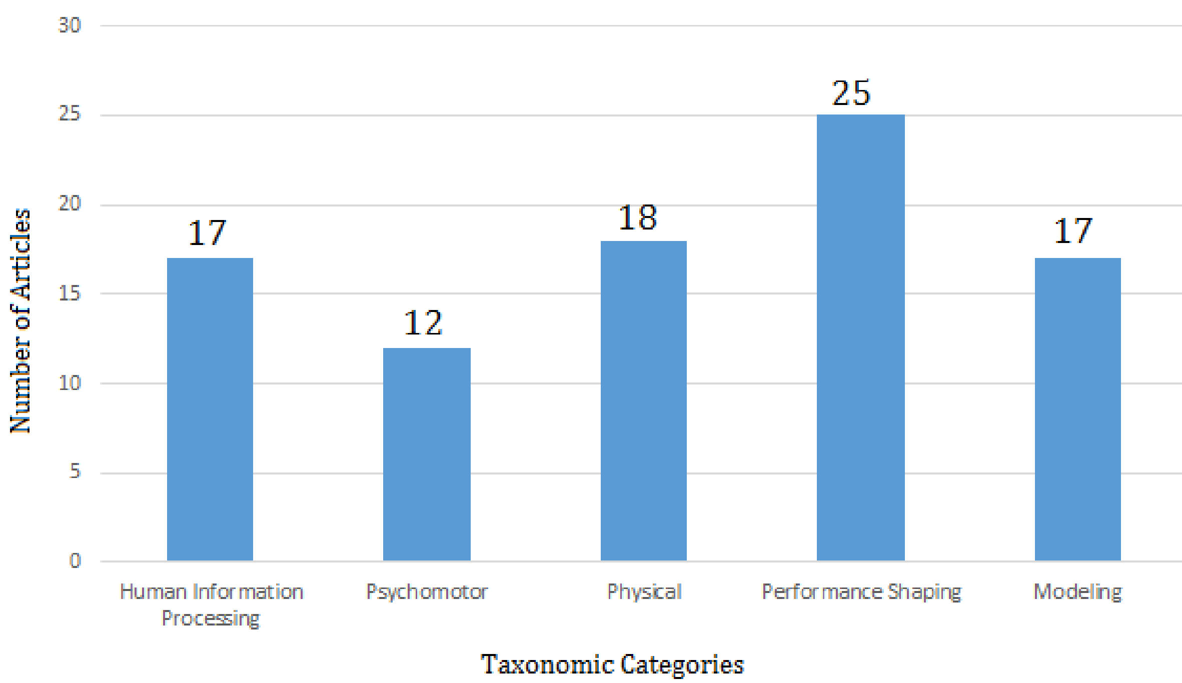


Figure 3: Cumulative Number of Publications on Human Temporal Variability

A summary of the articles by taxonomy category is provided in Table 2. The format of Table 2 provides the article authors as well as the relevant contributions of each article towards understanding temporal variability in human performance. In addition, the table also provides a normalized (Google Scholar Citations/Years in publication) indicator of each article's relative strength in contributing to the reviewed evolutionary research. The parenthetical number under each taxonomic heading identifies the number of studies reviewed for that category.

Table 2: Summary of Included Articles by Category

Taxonomic Topic	Authors	Key Contributions	# Citations (Google Scholar)/Years in Publication
Human information processing (17)	Altmann & Trafton (2007)	Examined the time-course recovery response following a cognitive task interruption. Sampled over 13,000 interruptions to obtain stable data. Data showed response times dropped in a smooth curvilinear pattern for the first 10 responses of post interruption performance. This indicated the ability of the cognitive system to retrieve displaced mental context from memory incrementally, with each retrieved element adding to the set of primes facilitating the next retrieval.	13.56
	Balota & Yap (2011)	Examined the influence of using mathematical functions to fit empirically derived response time distributions and plotting as a function of conditions	12.33
	Bustamante, Bliss & Anderson (2007)	Investigated the effects of varying the threshold of alarm systems and workload on human response to alarm signals and performance on a complex task. Results showed that participants responded significantly faster to true alarm signals when they were using the	2.89

		system with the highest threshold under low-workload conditions. Results also indicated that changing the threshold of the alarm system had a significant effect on overall performance and this effect was greater under high-workload conditions.	
	Kamienkowski & Sigman (2008)	Investigated the timing and characteristics of human response time variability for parallel sensory and motor operations.	1.22
	Lin, Jou, Yenn, Hsieh, & Yang (2009)	Examined the effectiveness of information presentation and task operation in a complex STS from the human information processing (HIP) perspective. Influences and implications to staffing are also discussed.	0.78
	Madden et al. (1999)	Used positron emission tomography (PET) to measure age-related changes in regional cerebral blood flow. Separate PET scans were conducted during Encoding, Baseline, and Retrieval conditions. The complete reaction time (RT) distributions in each task condition were characterized in terms of an ex-Gaussian model (convolution of exponential and Gaussian functions). The data suggest that the attentional demands of this task are relatively greater for older adults and consequently lead to the recruitment of additional neural systems during task performance.	8.78
	Martin (2009)	Developed a general theory of reaction time (RT) distributions in psycho- logical experiments derived from the distribution of the	0.78

		quotient of two normal random variables, task difficulty and the external evidence that becomes available to solve it.	
	Palmer, Horowitz, Torralba, & Wolfe (2011)	Examined multiple trials from different benchmark visual search tasks and evaluated the ability of four popular functions to capture the resulting empirical RT distributions.	6.67
	Proctor & Vu (2006)	Reviewed the history of human information processing and its relation to human factors and human–computer interaction (HCI).	3.00
	Pyy (2000)	Presented qualitative and quantitative findings of an evaluated method to study human reliability in decision situations related to complex socio-technical systems.	2.11
	Roth, Patterson, & Mumaw (2001)	Introduced basic concepts of cognitive engineering and used examples to illustrate common design pitfalls that have led to poor human-computer systems.	0.22
	Shi & Shi (2013)	Designed a universal method to produce various degrees of mental workload and explore its effect on driver reaction time through a driving simulator.	0.22
	Sugarman (2011)	Examined integration of human response time and data from predetermined time systems to carry out socio-technical system accident analysis.	0.22
	Teichner, Williams, Ekel, & Corso (1979)	Presented a comprehensive theory of human information processing along with four studies designed to test predictions based on the theory.	0.00

	Van Zandt (2002)	Examined response time analysis, including estimates and the means and variance, outlier techniques, estimation of distribution parameters, and function estimation. The use of distributional analysis in testing processing models is also discussed.	15.11
	VanRullen & Thorpe (2001)	Used dual task event related potential where targets of one task are intermixed among distractors of the other to show visual categorization involves different mechanisms with different time courses.	55.67
	Waters (1998)	Examined the influences of nicotine on human selective attention in a detailed fashion using the Garner speeded classification task.	0.89
Psychomotor (12)	Abrams, Meyer, & Kornblum (1990)	Examined the detailed nature of coordination between the eyes and limbs during movement production by the ocular and manual motor control systems.	10.38
	Bedny & Karwowski, (2013)	Conducted a study of positioning actions using a functional analysis approach of activity where the activity is considered a self-regulative system. Previous research studied positioning motor actions with two targets, this research considers not just two, but four targets. Results of this study created new data related to the properties of the regulation process for positioning actions and supports that both cognitive and emotional-motivational mechanisms of activity regulation	1.67

		are important factors in error analysis.	
	Bliss & Chancey (2014)	Trained participants to react to alarms using sensor activity patterns. Analyses revealed more appropriate and quicker reactions when participants were trained and when the alarms were reliable..	0.50
	Bootsma, Marteniuk, MacKenzie, & Zaal (1994)	Empirically investigated how size of an object to be grasped influences the time for a prehensile movement to be completed.	9.09
	Borah (1995)	Investigated the feasibility of using eye point of gaze and head control of a display cursor, in place of, or to supplement manual control for cursor positioning tasks.	0.38
	Dumont & Mazer (2013)	Obtained descriptive sample data for age groups from 5 to 10 years, identified factors associated with performance, and examined the inter-rater reliability, internal consistency and construct validity of the test in a sample of typically developing children.	0.67
	Eskenazi , Rotshtein, Grosjean, & Knoblich (2012)	Tested whether motor activation corresponds to the difficulty of the observed action, using Fitts' law. The results revealed activation in the motor system during action observation is not driven by perceptual parameters but by the motor difficulty of the observed action.	2.25
	Itami, Antonio, & Mendes (2015)	Analyzed the reaction times obtained from participants in a psychomotor activity with a large number of trials without breaks and investigated the learning in terms of average values and their respective variability. Results	0.00

		indicated that the learning can be associated with a scale factor acting over the reaction times.	
	Lin & Wu (2013)	Examined innate differences between touch screens and standard physical keypads in the context of numerical typing and elimination of confounding issues. Effects of precise visual feedback and urgency of numerical typing were also investigated. The results showed that touch screens were as accurate as physical keyboards, but reactions were indeed executed slowly on touch screens as signified by both pre-motor reaction time and reaction time.	0.33
	Lin, Radwin, & Vanderheiden (1992)	Used a Fitts' Law task to determine how control display gain influences performance with a head controlled input device and compared relative sensitivity between head control and hand/arm control.	2.54
	Miles & Proctor (2012)	Investigated the relationship between three of the most commonly used spatial stimulus modes, arrows, locations, and location words, using correlations of compatibility effects between each of the modes as well as compatibility effects at different segments of their response time distributions.	3.25
	Seow (2005)	Discussed the common information theoretical concepts of the Fitts' and Hick-Hyman Laws, and then examines each law with respect to its origins, theoretical formulation, theoretical development, research, and applications and examined the	7.82

		possible contributing factors responsible for the failure of the Hick-Hyman Law to gain momentum in the field.	
Physical (18)	Abbott, Liu, Chua, & Chang (2011)	Used work measurement methods for all aspects of a ships block's construction to develop a probabilistic model for the construction man-hours and to provide better estimates of the man-hour required to support planning and scheduling.	0.2
	Aft (2010)	Examined the value in having established work measurement time standards that are the output of the work measurement process. These standards affect every facet of an organization's operations and business functions.	0.166667
	Bedny, Karwowski, & Voskoboyniko v (2015)	Examined the behavioral components of work activity in time studies. Described insufficiency of method time measurement (MTM-1) system in analyzing the strategies for task performance and studying the logical organization of motor and cognitive actions.	0
	Bohannon & Andrews (2011)	Consolidated data from multiple studies to provide normative data that can serve as a standard against which individuals can be compared.	32.8
	Chen & Joyner (2006)	Simulated a military mounted environment and conducted experiments to examine the workload and performance of the combined position of gunner and robotic operator. Examined how individual difference factors such as perceived attentional control	0.2

		and spatial ability were related to the task performance.	
	Choodoung & Smutkupt (2012)	Performed a motion time study of the assembly process for wood joints with Methods Time Measurement-2 (MTM-2) and DFA (Design for Assembly). Completed an assessment of the ability to assemble in the feeding and fitting stages with LUCAS Assembly Evaluation Method	0
	Christmansson, Falck, Amprazis, Forsman, Rasmussen, & Kadefors (2000)	Evaluated use of a motion time study tool called ErgoSAM based on a higher-level method-time-measurement (MTM) system called SAM. The ErgoSAM method considers information on weight handled or forces applied and work zone. The method is designed to predict the physical demands of work postures, force, and repetition according to the Cube model.	2
	Collins & Kuo (2013)	Examined hundreds of over ground walking steps by healthy young adults (N = 14, age < 40 yrs.). Identified that slow fluctuations in self-selected walking speed could explain most variance in step length. Identified factors not related to balance which may reveal what aspects of walking are most critical for the nervous system to control.	4.333333
	Department of Defense (1997)	Redesigned the DoD Work Measurement/Labor Standard program to enhance performance and develop a general architecture for standardizing automated support for industrial engineering techniques.	0

		Combined the traditional work of developing labor standards and manpower standards/requirements with efforts of providing process improvement, economic analysis, quality programs and organizational improvements. Developed data on required time and manpower to perform identified work, assisted managers in tracking results and performing variance analysis of expectations compared to actual results, provided work analysis and continual improvement assistance.	
	Harman, Frykman, Pandorf, Tharion, Mello, Obusek, & Kirk (1999)	Experimentally evaluated the physiological, biomechanical, and maximal performance response of soldiers carrying light, medium, and heavy loads.	1.411765
	Karwowski (2013)	Presented focused considerations for the development of an ideal human observer concept. Examined manpower performance components and human error and reliability estimates to use in calculating operator error values for any functional, tactical, or operational task.	0
	Kuhlang, Edtmayr, & Sihn (2011)	Introduced a methodical approach to connect Value Stream Mapping (VSM) and Methods-Time Measurement (MTM) to offer advantages in reducing lead time and increasing productivity based on lean principles and standardized processes.	10.6
	Laring, Forsman, Kadefors, &	Developed an ergonomic complement to the modern MTM system called SAM to give insight into the future ergonomic quality	5.285714

	Örtengren (2002)	of a planned production. Identified a method that relies on two additional pieces of information to the analysis: the zone relative to the operator's body in which the movement takes place or ends, and the weight or force involved in the operation.	
	Nakayama, Nakayama, & Nakayama (2001)	Proposed a method for setting standard time using a work achievement quotient approach. This method is intended for work measurement of a small manufacturing volume or of a long cycle time where conventional methods such as Predetermined Time System (PTS) such as the Work Factor (WF) and Method-Time Measurement (MTM) may not be practical.	0.533333
	Razmi & Shakhs-Niyaee (2008)	Developed a tailored predetermined time study method using special time tables developed by the combination of MOST and work time table standards.	1.25
	Sabaric, Brnada, & Kovacevic (2013)	Presented general features of the MODAPTS (Modular Arrangement of Predetermined Time Standards) method and its application in the warping process during making fabrics.	0
	Yadav (2013)	Presented a Knowledge Based Design Methodology (KBDM) for automated and manual assembly lines measurement with help of Maynard operating sequencing technique (MOST). This method can be applied equally well to single, multi- and mixed-product assembly lines with either	0.333333

		deterministic operation times or stochastic operation times.	
	Yogev-Seligmann, Hausdorff, & Giladi, (2008)	Examined the role of executive function and attention in healthy walking and gait disorders while summarizing relevant literature. We Described the variety of gait disorders that may be associated with different aspects of executive function, and discuss the changes occurring in executive function as a result of aging and disease as well the potential impact of these changes on gait.	83.25
Performance Shaping (25)	Andreassi & Huntley (1967)	Vigilance performance and physiological responses with variable interval (VI) and fixed interval (FI) signal patterns were studied. Reaction time (RT) was used as the performance measure while heart rate (HR), palmar skin conductance (PSC) and galvanic skin responses (GSRs) were the physiological measures. Results indicated that there was a tendency for RT to be faster with fixed interval as compared to variable interval.	0.06
	Blackman, Gertman, & Boring (2008)	Presented cognitively based human reliability analysis quantification technique with the intent to develop a defensible method that would consider all factors that may influence performance.	4.00
	Carey & Kacmar (1997)	Examined the impact of technology on a number of factors including time to complete task, member satisfaction, perceived information load, the number of	4.89

		contributing transactions, and task complexity.	
	Chan, Shum, Law, & Hui (2003)	Studied the precise effect of control knob position, indicator type and scale side on strength of stereotype, index of reversibility (IR) and response time for a horizontal display/rotary control arrangement.	1.46
	Dietz, Weaver, Sierra, Bedwell, Salas, Fiore, Smith-Jentsch, & Driskell, (2010)	Presented an innovative theoretical approach for unpacking the temporal and interactive effects among performance stressors forming a foundation for understanding their impact on dynamic episodes of individual and team performance.	0.17
	Dixon, Wickens, & Chang (2005)	Studied the performance of licensed pilots in flying both single-UAV and dual-UAV simulated military missions. Practical implications for the study include the suggestion that reliable automation can help alleviate task interference and reduce workload, thereby allowing pilots to better handle concurrent tasks during single- and multiple-UAV flight control.	8.91
	Dunn & Williamson (2012)	Studied forty participants completing one of two computer-based tasks differing in terms of cognitive complexity along with scales rating workload, boredom proneness, fatigue, and task characteristics. Results indicate similar levels of subjective fatigue between tasks with no difference in fatigue ratings between the tests. Performance tests however showed that simple choice reaction time task indicated clear evidence	0.75

		of the influence of time on test as response times and errors increased with task duration.	
	Fereidunian, Zamani, Fatah, Lesani, Lucas, Kharzami, & Torabi (2010)	Investigated the relationship between human-automation systems and the factors which shape their performance. Ranked the most influential Performance Shaping Factors in order of their influence in a practical automation system.	0.17
	Grabbe & Allen, (2012)	Examined the effects of stimulus-stimulus and response-response cross-task compatibility and aging on dual-task performance.	0.50
	Griffith, & Mahadevan (2011)	Discussed the importance of the effects of fatigue on performance, the difficulties associated with defining and measuring fatigue, and the current status of inclusion of fatigue in human reliability methods.	5.20
	Hart & Staveland (1988)	Identified the results of a multi-year research program studying the factors associated with variations in subjective workload within and between different types of tasks. Task-, behavior-, and subject-related correlates of subjective workload experiences varied as a function of difficulty manipulations within experiments, different sources of workload between experiments, and individual differences in workload definition.	201.43
	Hocking, Silberstein, Lau, Stough, & Roberts (2001)	Administered a select range of psychometric tests and imaged functional brain electrical activity to investigate the impact of thermal stress on cognitive	5.07

		performance including cognitive time variability.	
	Koh, Park, & Wickens (2014)	Examined differences on task management behaviors between differing levels of experience, and correlated indices of task management with levels of performance evaluated by subject matter experts.	1.50
	Lee, Kim, Ha, & Seong (2011)	Derived performance shaping factors (PSFs) and a new qualitative evaluation framework for these PSFs. The PSFs from various methods are collected and grouped into categories, and then human factor (HF) issues are analyzed and derived to be used as an evaluation framework for PSFs.	1.40
	Mackieh & Cilingir (1998)	Examined the effects of motor variables, decision-making mechanism, complexity of information presented, intelligence levels, and emotional states of subjects on human performance.	0.72
	Marras & Hancock (2014)	Examined advancing the level of sophistication in the practice of human factors and ergonomics to begin considering the totality of the human-system behavior and performance in combination with systems design.	5.00
	Marusich, Buchler, & Bakdash (2014)	Investigated how varying levels of available information affects human decision-making. Findings raise questions about human capabilities for information fusion given the high volume of information in military networks. Results also suggest that decision support systems may enhance human capabilities for fusing and disambiguating information.	1.50

	Matthews, Warm, Dember, Mizoguchi, & Smith (2001)	Studied the effects of naturally-occurring colds on visual attention, psychomotor performance and subjective indices of stress. Affective, motivational and cognitive stress state dimensions were measured. Results indicated a direct effect of colds on simple reaction time, whereas the cold effect on vigilance appeared to be statistically mediated by reduced task engagement.	1.53
	Micalizzi & Wickens (1981)	Described the selective assessment of primary task workload, within the framework of a multiple resources model of human information processing. Performed reaction time tasks alone and concurrently with a primary task of interest.	0.06
	Mracek, Arsenault, Day, Hardy, & Terry (2014)	Demonstrated a longitudinal, multilevel approach to examine the dynamic relationship between subjective workload and performance over a given period of activity involving shifts in task demand. Results showed that both between- and within-person effects were dynamic. Higher subjective workload reflected performance problems, especially more downstream from increases in task demand.	1.50
	Murthy & Kerr (2003)	Investigated the interaction between communication process goals and communication modes. Results revealed a significant interaction between communication mode and communication process goals.	6.62

	Podofillini, Park, & Dang (2013)	Applied a task complexity measure to quantify the complexity of procedure-guided tasks, and evaluate task complexity issues relevant to human reliability analysis methods.	2.67
	Stern & Brown (2005)	Developed and used a test task which incorporates the need for visual search activity as well as involves a cognitive component. For non-sleep deprived subjects, this task demonstrates lapses in performance as indicated by significant changes in reaction time (RT) as a function of Time-on-Task.	0.00
	Weaver, Foxe, Shpaner, & Wylie (2006)	Assessed the effect that unexpected task constraint, following self-generated task choice, has on task switching performance. Results suggested that when participants choose to switch tasks, they prepare for that switch in anticipation of the stimulus, and the preparation is durable such that it cannot be undone readily without an associated time cost.	0.20
	Young, Brookhuis, Wickens, & Hancock, (2015)	Provided a general overview of the current state of affairs regarding the understanding, measurement and application of mental workload in the design of complex systems over the last three decades. Concludes by discussing contemporary challenges for applied research, such as the interaction between cognitive workload and physical workload, and the quantification of workload 'redlines' which specify when	22.00

		operators are approaching or exceeding their performance tolerances.	
Modeling (17)	Allender (2000)	Discussed the U.S. Army's push toward use of simple equations, stochastic task network modeling, or representation in force-on-force models in simulation based acquisition. Provided underlying rationale and examples of models developed with the capabilities present in IMPRINT (the Improved Performance Research Integration Tool), developed by the Human Research and Engineering Directorate of the U.S. Army Research Laboratory.	1.81
	Azarkhil & Mosleh (2014)	Developed a method to explicitly model the operating crew of complex systems as an interactive social unit and investigated the dynamic behavior of the team under upset situations through a simulation method. An object based modeling methodology is applied to represent system elements and different roles and behaviors of the members of the operating team.	0.00
	Badler, Albeck, Lee, Rabbitz, Broderick, & Mulkern (2005)	Presented a next generation digital human modeling test-bed that includes a scriptable interface, real-time collision avoidance reach, empirical joint motion models, a versatile locomotion engine, motion capture, and synthetic motion blends and combinations.	2.00
	Chang & Mosleh (2007)	Discussed the Information Decision and Action in Crew (IDAC) context for human reliability analysis (HRA) and	16.33

		example applications. The model is developed to probabilistically predict the responses of the control room operating crew in nuclear power plants during an accident, to include temporal variability for use in probabilistic risk assessments (PRA). The operator response spectrum includes cognitive, emotional, and physical activities during the course of an accident.	
	Department of Defense (1989)	Objective is to acquaint potential users with a broad range of models that may be used to predict aspects of human performance during the system development process.	1.48
	Fuller, Reed, & Liu (2010)	Addressed the divide between cognitive and physical human models by integrating a cognitive human model with a physical human model. This new combined model used the advantages of each type of model to overcome the weaknesses of the other. The capabilities of the integrated model are evaluated in terms of modeling a task scenario with both cognitive and physical components: driving while performing a secondary in-vehicle task.	1.00
	Gregoriades & Sutcliffe (2008)	Described a method and a tool for analyzing and predicting workload for the design and reliability of complex socio- technical systems. Concentrated on the need to assess workload early in the design phase to prevent systems failures. The method is supported by a tool that enables scenario-based validation of prospective socio-technical	4.25

		systems designs such as command and control rooms of military vessels. The approach combines probabilistic measures of human performance with subjective estimates of workload.	
	John, Patton, Gray, & Morrison (2012)	Described a combination of theory based tools to estimate the time variability of skilled human performance in real-time, safety-critical tasks. Discussions covered the tools, their integration, and provided a concrete example of their use.	0.75
	Kuo & Wang (2009)	Used the Method Time Measurement (MTM) system as the basis for defining the operational motion semantics to generate human motions in a digital environment. By using the MTM semantics as the motion command and applying simple rules for locomotion, the upper and lower limb motions and the gesture of a Digital Human Model can be generated automatically. The virtual simulation results obtained from this developed system can be used to evaluate job and workplace design, as well as conduct ergonomic evaluations.	2.57
	Li (2013)	Introduced a methodology for modeling and simulating nuclear power plant operators' knowledge-based behavior and further demonstrated that it is possible to model individual operator's underlying cognitive processes and generate realistic response scenarios through dynamic simulation.	2.33

	Mason, Baines, Kay, & Ladbrook (2005)	Explored the use of probability density functions to represent the variation of worker activity times within discrete event simulation (DES) models.	1.64
	Reer (1994)	Presented a probabilistic method for analyzing human reliability under emergency conditions. The model used time window and organization input data, and enabled a quantitative comparison between several organization alternatives for an emergency time constrained response. The method subdivides the whole emergency procedure into single steps and results in relatively high accuracy performance time distribution assessment.	0.45
	Rouder, Lu, Speckman, Sun, & Jiang (2005)	Presented a statistical model for inference with response time (RT) distributions. The hierarchical model provided a means of estimating the shape, scale, and location (shift) of RT distributions as well as between-subjects and within-subjects variability.	13.09
	Stanton & Baber (2008)	Applied two modeling approaches to the same problem to see if they arrived at the same conclusion. The first modelling approach used the alarm initiated activity (AIA) model. This approach is useful for indicating general response times in emergency events, but it cannot comment in detail on any specific case. The second modelling approach employed a multi-modal critical path analysis (CPA) technique. This research has application to the modelling of human responses to emergency	2.38

		events in all domains and could be used in a predictive manner to anticipate how long human operators of safety-critical systems might take to respond in emergency scenarios.	
	Stary & Peschl (1998)	Examines epistemological and methodological assumptions in the field of cognitive modeling as well as their implications for user interface design	1.28
	Ulrich & Miller (1993)	Described response time mechanisms that could generate the log normal distribution and showed how specific response time models can be constructed within the framework.	3.74
	Wong, Walters, & Fairey (2010)	Focused research on the modelling method developed while creating a Discrete Event Simulation (DES) model for the Rendezvous, Proximity Operations, and Docking (RPOD) phase for the International Space Station.	0.50

2.3.1.4 Literature Review Validity Risk and Mitigation

A main threat to the validity of this summary literature review is lack of completeness. The risk of this threat depends on the selected keywords and the limitations of the employed search engines. To decrease the risk of an incomplete keyword list, well-known and heavily cited articles were used to develop the initial key word search list. An iterative and evolutionary approach was then used to build subsequent keyword search lists until the majority of the identified articles were duplicates of previous searches. New keywords were added when the keyword list was not viable in producing new and unique representative literature. In order to omit the limitations imposed as the result of employing a particular search engine, multiple search engines were used. A second significant threat to the validity of this literature review is the sheer volume of the articles identified relative to the time and attention span available for the review. As the review of the literature itself indicates, the repetitive nature of reviewing hundreds of articles for content under perceived and/or real time constraints increased the likelihood of erroneously eliminating relevant articles. Mitigation of this risk was accomplished by parsing the articles into manageable, discrete quantities of 20-40 articles depending on article length and implementing sufficient time separation between reviews. Another important concern is assuring the rigor and robust nature of the article segregation taxonomy to assure proper classification and analysis of the papers. To avoid a taxonomy with insufficient breadth to classify the selected papers, an iterative content analysis method was continuously used to assure sufficient taxonomy categories for every new concept encountered in the literature review.

2.3.2 Mechanisms of Human Task Performance

Successful human task performance requires the cumulative effort of the cognitive and physical abilities of the individual completing the task. In this context, *cognitive abilities* refers to the ability to process information and make task decisions based on environmental perceptions, knowledge, and memory. *Physical abilities* describes the coordinated muscle action needed to complete sustained, effortful, muscular work.

In complex, adaptive systems where humans are central to the system, they will respond, as enabled by their attention, to surrounding stimuli based on how they perceive their environment through visual and auditory perception (Salvendy, 1997). This response, dependent on the type of action required, will necessitate that individuals leverage their cognitive and physical abilities in completing the task. These actions occur principally in series with one another, although some components of the response may be conducted in parallel.

The task environment, on the other hand, consists of all the elements, both internal and external to the individual, that impact the human response within the STS context (Marras & Hancock, 2014). A generalized view of the STS task performance relationship is provided in Figure 4.

Figure 4 identifies that, for a given STS, interdependencies between the defined performance environment and the human component of the system impact both the physical and cognitive performance characteristics of the individual. In addition, the dotted relationship recognizes existing relationships and likely interplay between the physical and cognitive abilities of the individual in completing the task (Marras & Hancock, 2014). An outcome of these relationships is the cumulative nature of the time required for both the cognitive and physical functions in the completion of a task.

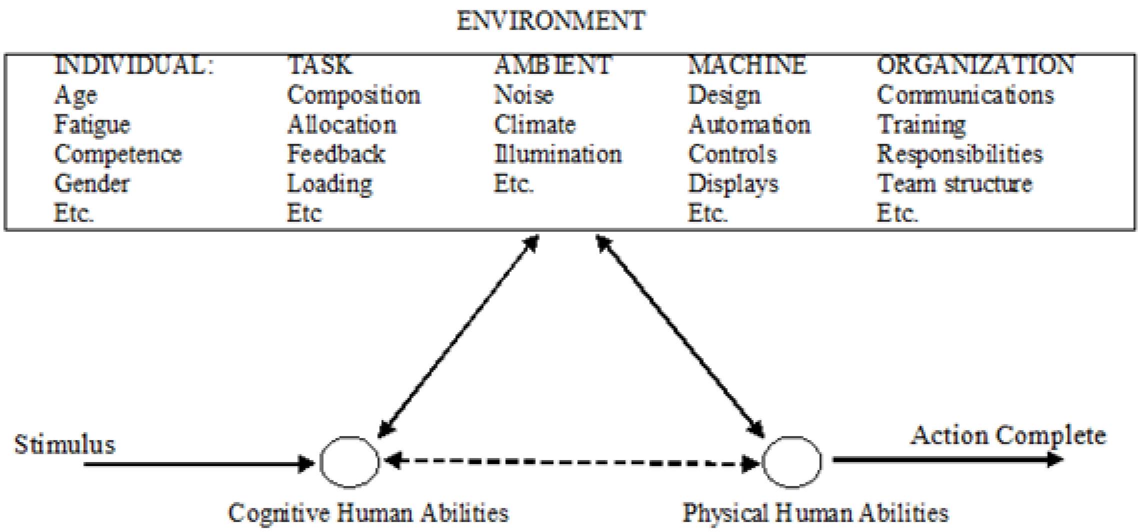


Figure 4: Relationships and Interdependencies of Physical and Cognitive Capacities (Adapted from Marras & Hancock, 2014)

2.3.3 Taxonomy of Human Task Abilities

Ability is the competence an individual has in the performance of an activity or occupation because of their skill, training, or other qualification (Ability[Def 2], n.d.). Empirically, it is seen that considerable differences in ability do exist between workers. Differences in inherent knowledge, physical capacity, health, trade knowledge, physical dexterity, and training can cause one operator to consistently outperform another (Freivalds, 2009). In this sense, abilities are part of the traits that affect an individual's capability to become skillful, under the influence of external factors, when learning a new task. Therefore, an effective taxonomy of human abilities recognizes the premise that there exists a definable set, both cognitive and motor, that can be used in the performance of a task. Ability taxonomies have been the subject of considerable research and debate over the past several decades, with many variations both proposed and employed. In fact, as early as 1938, Thurstone created a rudimentary taxonomy that viewed intelligence as composed of a small set of primary mental abilities that include verbal comprehension, numerical reasoning, word fluency, and memory. In his studies, Guilford (1956, 1959) suggested taxonomies of intellectual capabilities that described tasks using an information-processing model. More recently, Fleishman and Quaintance (1984) identified a taxonomy that could be used to characterize the abilities required in most performance situations. This taxonomy consisted of 52 human abilities ranging from verbal comprehension to speed of limb movement. While Fleishman's taxonomy is certainly one of the most comprehensive, the breadth of abilities he considers often exceeds those typically used by researchers in the field of Human Factors and Ergonomic Sciences. For example, in his work with the U.S. Army to predict crew performance degradation as a function of influencing factors, Roth (1992) developed a taxonomy based on the premise that the tasks humans perform can be broken down into a basic set of core abilities. His taxonomy consisted of five skill types as described in Table 3.

Table 3: Taxonomy of Required Task Abilities

Ability	Description
Attention	The ability to attend actively to a complex stimulus for extended periods of time in order to detect specified changes or classes of changes that indicate the occurrence of some phenomenon that is critical to task performance
Perception	The ability to detect and categorize specific stimulus patterns embedded in a complex stimulus
Cognitive	The ability to apply concepts and rules to information from the environment and from memory in order to select or generate a course of action or a plan. This includes communicating the course of action or plan to others.
Psychomotor	The ability to maintain one or more characteristics of a situation within a set of defined conditions over a period of time, either by direct manipulation, or by manipulating controls that cause changes in the characteristics
Physical	The ability to accomplish sustained, effortful muscular work.

(Proposed by Roth, 1992)

Consistent with the direction taken by Roth (1992), most human factors practitioners prefer to select an ability taxonomy that adequately covers the range of skills needed for the task but is sufficiently discriminating to provide a manageable number of categories for an analyst to use. While selection of the best taxonomy typically depends on the particular tasks and stressors, the "case-by-case" approach tends to create a disparate and disjointed application of inconsistent taxonomies to similar tasks.

2.3.4 Taxonomy of Human Performance Influencing Factors

The factors influencing time variability in human task performance play a pivotal role in every STS. As a result, understanding their impacts is a necessity for complex STS environments where adverse time impacts could result in catastrophic consequences. This necessity has motivated a prodigious amount of far-reaching and sometimes exhaustive research across a wide breadth of STS applications, including the nuclear power industry, military battle-space operations, railway activities, manned space flight, and marine/offshore applications. This research is typically directed toward the more holistic analysis of Human Reliability Analysis (HRA) and requires integration of the full range of human abilities, from environmental information processing through performance of all necessary physical activities. In addition to the identified critical STS applications, industrial sectors such as manufacturing have also started to integrate the impacts of factors affecting human performance into their operations as a means of reducing cost and improving quality (Bubb, 2005).

As mentioned earlier, successful task completion within a prescriptive time standard associated with the task is often a critical sub-element of HRA. As a result, human temporal variability in task performance and HRA are inextricably linked. Concepts and traits associated with development of Performance Influencing Factors (PIFs) for HRA apply equally well and tend to match factors affecting time variability in human task performance. Subsequently, discussion and approaches found within literature in assessing human temporal variability are consistent with those methods and outcomes used in developing PIFs for human reliability analysis.

PIFs, also called Performance Shaping Factors (PSF) in HRA, are a heavily researched area and accepted standard within the field of human performance analysis. Similar to the discussions on categorical taxonomies for human abilities, PIF taxonomies have been the subject of considerable research and academic debate over the past several decades. To date, a consensus on which PIFs should be used and the appropriate number of PIFs to include in a method or analysis has yet to be reached (Boring, 2010). Taxonomies using a single PIF as well as some that use upwards of 200 PIFs have been addressed in literature. They are often developed and tailored in unique and niche frameworks to assess task performance variations inherent to specific human conditions, as well as in discrete factors that influence the performance of an individual or team in a specific environment (Kim & Jung, 2003; Mindock & Klaus, 2014). The most studied influences are those that result in negative outcomes, although positive impacts on human performance can and do occur.

The research community has identified two principle types of PIFs: external and internal. In their work on human task performance influencing factors, Kunihide Sasou and James Reason (1999) define the task environment PIFs as either internal or external, where:

1. External PIFs are shared by people working within the same environment.
2. Internal PIFs are dependent, at least in part, on the individual.

Although internal PIFs are not necessarily independent of external PIFs, individuals may respond differently to the same external impact. Therefore internal PIFs are considered separately from external ones. Sasou and Reason (1999) also point out that most human work is performed by teams rather than individuals, particularly in complex human-system related processes, such as naval shipboard applications, nuclear power generation, commercial aviation, and the like. In these team configurations, Sasou and Reason define team PIFs as "factors arising from a group of people working together on a common project or task. They include lack of communication, inappropriate task allocation, excessive authority gradient, over-trusting, etc." (1999). However, the segregation effort of PIFs between external, internal, and team provides only an initial framework for analysis. In-depth analysis of the role of PIFs in human performance requires a detailed classification mechanism to quantify their role.

For example, Groth and Mosleh (2012) in their work have articulated a pragmatic classification approach in developing a set of fundamental principles to serve as guidelines for development of PIF sets and for expansion of proposed PIF hierarchal structures to more detailed levels. Specifically, their work identified that:

- Analysis should consider only those PIFs that directly impact the individual's performance
- Events must be parsed into sub-events consistently based on established rules
- PIFs must be defined orthogonally; i.e., they must be separately defined entities
- PIFs should be value neutral with the ability to expand in characterizing context

This PIF taxonomy approach as implemented by Groth and Mosleh (2012) is representative of the methods used by others in this research area, and provides a PIF hierarchal framework that is easily adapted to provide a framework for assessing time variability in human task performance. It can be expanded or collapsed accordingly by simply adjusting rows and/or columns to tailor specific task environment factors as shown in Table 4.

Table 4: PIF Taxonomy Table

Organization	**Team based**	**Individual**	**Ambient**	**Task**
Training Program -Availability -Quality	Communication -Availability -Quality	Attention	External Environment	Human System Interface
Corrective Action Program -Availability -Quality	Direct Supervision	Physical and Psychological Ability	Conditioning Events	System Response
Other programs -Availability -Quality	Team Coordination	Knowledge/ Experience	Task/Time Load	
Safety Culture	Team Cohesion	Skills	Other Loads	
Management Activities -Staffing -Scheduling	Role Awareness	Bias	Task Complexity	
Workplace Adequacy		Familiarity w/Situation	Stress	
Resources (procedures, tools, information)		Morale/Motivation /Attitude	Perception	

(Adapted from Groth and Mosleh, 2012)

Today, as a result of studies similar to those produced by Groth and Mosleh (2012) across a range of industries (Fereidunian, et al., 2010; Lee, Kim, Ha, & Seong, 2011; El-Ladan & Turan, 2012) there are more than a dozen PIF based methods in use without a consistent set of standard PIFs among the methods. This variability is understandable given that each set of PIFs is a reflection of the multitude of factors that can influence human performance, the different approaches used to distill them, and the various applications for which each is applied. However, it is also widely recognized that this inconsistency limits the utility of these methods and tools, expends considerable energy in creating detailed representations, and negatively impacts their application to more generic, complex, industrial or operational environments (Boring, 2010). Subsequently, it is generally accepted that continued research and consensus development in the formation of a standard vocabulary and structure for factors affecting temporal variability and analysis is warranted.

2.3.5 Human Task Performance: Temporal Response and Impacts

Once developed, good taxonomies provide the basis and conceptual framework for discussion and analysis of the time impacts associated with human task performance. They also support development of representative distributions resulting from the variance in human abilities and task environments. As a framework for discussion of human abilities, a taxonomy leveraging the work of Roth (1992) described above provides suitable context to cover the necessary range of impacts and distributions applicable to human temporal variability. Inclusion of additional abilities into a similar taxonomy would require only adjudication as to the applicable impacts and distributions. Similar suitability exists for application of the PIF taxonomy as adapted from Groth and Mosleh (2012) in discussing the induced temporal variability as a result of internal and external factors within the task environment. These patterns of applicability are tailored and employed throughout literature to identify and assess time impacts and distributions associated with inter- and intra-individual characteristics.

2.3.5.1 Attention, Perception, Cognitive: Human Information Processing Temporal Response

As described by Roth's taxonomy and in most generalized Human Factors literature, interplay between the independent elements of Attention, Perception, and Cognitive abilities allows them to be interwoven into a single human information processing model. Owing to the considerable research in this area over the years, multiple distinct approaches to human information processing have been developed (Salvendy, 1997). For this research effort, a heavily cited cognitive engineering framework as defined by Wickens (1992) was used to highlight the role these abilities play in human response time analysis. In this framework, human cognitive response time is a function of an operator's selective processing in the environment (attention), and their ability to provide the information received from the environment (perception) with some meaningful interpretation (cognitive ability). Response times developed under this context represent the time difference between the onset of a sensory stimulus and subsequent physical response. The literature shows that research in these areas has seen both a focus shift and significant acceleration over the last thirty years as a result of society's moving from an industrial, machine-driven environment to a complex, semi-automated environment dominated by computers. Despite this research scrutiny, or perhaps as a result of it, response times associated with human information processing remain one of the most challenging factors to clearly evaluate. This difficulty is due, at least in part, to the wide variety of analytical approaches and statistical methodologies that have been, and continue to be, used in assessing these response times.

Response time (RT) as an innate human ability has been an important measure in the investigation of human information processing and cognitive studies for well over the last century. In general, and until relatively recently, researchers have conducted the bulk of their quantitative analysis based on statistics associated with the mean, using analysis of variance

(ANOVA) and other similar methods (Van Zandt T. , 2002). However, as the result of sustained research and publishing (Luce, 1986; Heathcote, Popiel, & Mewhort, 1991; Balota & Spieler, 1999; Olivier & Norberg, 2010), it is now generally accepted that variability of human response time is not accurately represented by a normal distribution (O'Boyle & Aguinis, 2012). Specifically, it is now recognized within the Human Factors and Ergonomics community that human RT distributions are virtually always positively skewed, with RTs clustering at the faster end of the scale (Balota & Yap, 2011). As such, significant research has been devoted to the understanding of the best distributions to describe response times. This research has resulted in a broad array of sophisticated and ad-hoc distributions including log-normal (Ulrich & Miller, 1993; Reer, 1994; van der Linden, 2006), ex-Gaussian (Matzke & Wagenmakers, 2009; Olivier & Norberg, 2010), and Weibull distribution applications, to name a few dominating current literature. In his work, Martin (2009) created a useful classification table of common human response time distributions. As adapted below in Table 5, Martin's (2009) classifications clearly illustrate the broad application of tailored distributions currently used by Human Factors researchers.

Table 5: Classification of Applied Response Time Distributions

Type	**Distribution**	**Dominant Term**	**Shape (Log Scale)**	**Shape (Log-Log)**
Exponential	Exponential Gamma Inverse-Gaussian Ex-Gaussian Ex-Wald	$e^{-\lambda t}, \lambda > 0$	Linear Decrease	Exponential Decrease (slow)
Quadratic-Exponential	Normal	e^{-kt^2}	Quadratic Decrease	Exponential Decrease (fast)
Lognormal	Lognormal	$\frac{1}{t} e^{-(\log t)^2}$	Quasi-linear Decrease	Quadratic Decrease
Power-Law	Pareto Cauchy Recinormal Fielers	$t^{-\alpha}, \alpha > 1$	Logarithmic Decrease (from t_{min})	Linear Decrease (from t_{min})
Power Law w/cutoff	DDM-Small	$t^{-\alpha} e^{-\lambda t}$ $\alpha > 1, \lambda > 0$	Power law until t_{max} and linear from t_{max}	Power law until t_{max} and exponential from t_{max}
Stretched-Exponential	Weibull	$t^{\beta-1} e^{-\lambda t^{\beta}}$ $\lambda, \beta > 0$	Above-linear decrease	Below-linear decrease

(Adapted from Martin, 2009)

Empirical distribution mapping and model development are not the only methods found in literature analyzing temporal components of human information processing. Human Cognitive Reliability (HCR) is another method characterized in current research. This research relies on the application of the skill-rule-knowledge (SRK) model proposed by Rasmussen (1984). The framework for this model is based on human behavior being parsed into skill-, knowledge-, and rule-based components relative to the human information processing level used (Di Pasquale, 2013). Under this model, normalized time reliability curves, approximated by three parameter

Weibull distributions, are derived using simulator data and small-scale tests. These curves are then plotted using the three parameter categories (skill, rule, knowledge) of human information processing as shown in Figure 5.

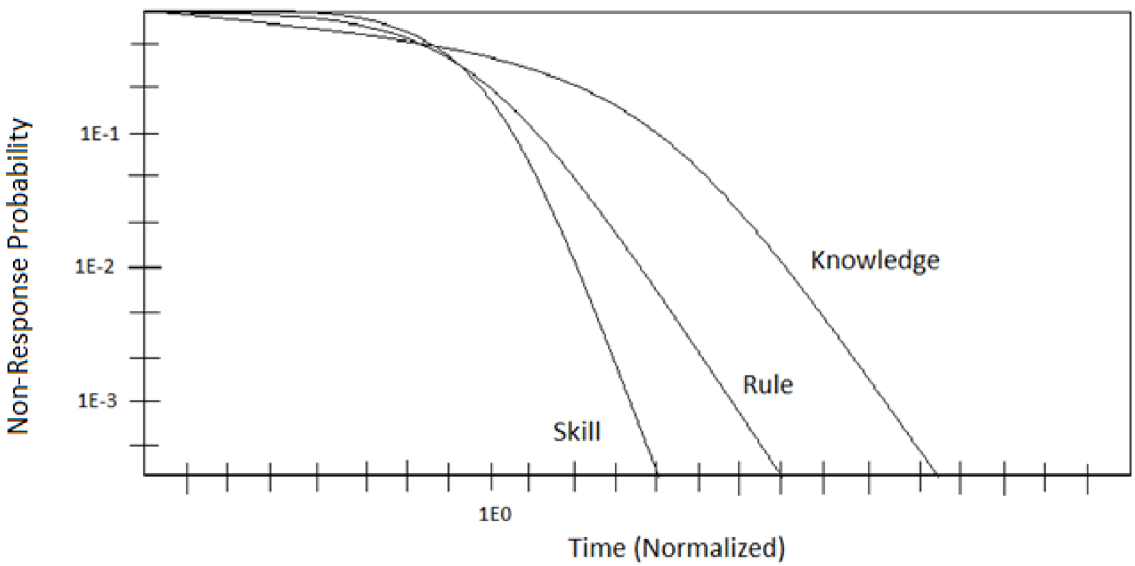


Figure 5: Rasmussen’s Skill-Knowledge Rule Model
(Adapted from Hannaman & Spurgin, 1984

Once developed, this model accounts for the variability in human response based on the likelihood, or probability, that a response will occur within a given time and inherently takes into account the information based hierarchal nature of the individual elements of Roth’s (1992) taxonomy.

2.3.5.2 Psychomotor Ability: Temporal Response

Psychomotor activity is typically the direct response of cognitive activity and results in muscle commands and mediated muscle activity. Psychomotor ability specifically refers to the skilled performance of physical functions, reflex actions, and interpretive movements. Traditionally, these types of objectives are demonstrated by physical skills such as movement, coordination, manipulation, dexterity, grace, strength, and speed-actions that demonstrate fine motor skills such as those studied in the field of Human Computer Interactions (HCI). In fact, over the last three decades, it is in the realm of HCI that significant momentum has emerged in the study of psychomotor temporal variability and its role in human task performance. Humans, in the performance of their task completion efforts, routinely interact with computers and need to negotiate the fine motor skills of quickly and accurately interfacing with complex STS interfaces.

Current research on psychomotor abilities and their role in human task performance typically rests on the foundation of Information Theory as developed by Shannon and Weaver (1949), applied by Fitts (1951,1954), Hick (1952, 1953), and Hyman (1953), and advocated for by Card, Moran, and Newell (1986). The result of the work by Hick and Hyman is a theoretical model used for reaction times in an environment of choice and is representative of the ambient task factors influencing human temporal variability. The work by Fitts is directly applicable to psychomotor response time and led to what has become known as Fitts' law as given in Fitts and Peterson (1964).

Fitts' law as originally presented recognizes that the human body has a limited capacity to transmit information in organizing motor behavior (Seow, 2005). As a result, Movement Time (MT), or task performance time, is proportional to the amount of information, in the form of feedback, required for producing the movement (Beamish, et al, 2009) as given by:

$$MT = a + b(ID) \tag{1}$$

In this equation, a and b are empirically derived model parameters and ID is the index of difficulty as defined by:

$$ID = log_2\left(\frac{2D}{W}\right) \tag{2}$$

where, D is the distance from the starting point to the center of the target, and W is the diameter of the target measured along the axis of motion.

A typical Fitts' Law evaluation plots execution times against IDs and shows with a regression test that there is a linear dependency. Fitts' Law has proven, through multiple studies, to be extremely robust in accounting for the variance in movement time performance under a variety of conditions (Smith, Henning, Wade, & Fisher, 2014). In addition, it has also proven to be applicable to a wide range of movements including prehension movements, mouse cursor movements, rotational movements, and foot movements (Lin, Radwin, & Vanderheiden, 1992; Grosjean, Shiffrar, & Knoblich, 2007; Bootsma, Marteniuk, MacKenzie, & Zaal, 1994; Abrams, Meyer, & Kornblum, 1990). However, Fitts' law does not always prove to be an accurate descriptor of movement or task performance time in all cases (Gan & Hoffman, 1988; Cha & Myung, 2010; Song, Clawson, & Radu, 2012) and, as a result, modified forms of Fitts' law have emerged. The two most commonly used variants of Fitts' law are those used by the HCI

(Human-Computer Interaction) community, called the Shannon formulation, and those based on the original Fitts model as used by most other researchers in the fields of ergonomics, engineering, and psychology (Hoffmann, 2013). Each one of these variants reflects individualized modification to more accurately model the time element of movement or task performance in its respective domains, however, both produce linear regression plots that are very similar to the one resulting from application of Fitts' Law. Consistent with the recurring theme in the evolving study of human task performance, an accepted or generalized standard has not yet emerged in the evaluation of psychomotor abilities and continues to be the subject of considerable discussion across the HF/E and HCI communities (Drewes, 2010; Hoffmann, 2013).

2.3.5.3 Physical Ability: Temporal Response

Physical ability as defined by Roth's (1992) taxonomy refers to the ability to accomplish sustained, effortful, muscular work in performing the motor actions required to complete the task. In the sense of actual task performance, it can be visualized as representing actions, postures, and motions used to complete the required activities. Physical abilities can vary on both an inter- and intra-individual basis, dependent on many factors such as gender, age, health and well-being, physical size and strength, aptitude, job satisfaction, and motivation, to name but a few (Freivalds, 2009). Output from the application of physical ability is typically evidenced in the form of human work, which can be measured through a variety of means and evaluated for a multitude of applications. The measurement of time in the performance of discrete physical activities is an often applied standard and is used to evaluate many different applications. In complex STS, these parsed time measurements can then be aggregated to define the nominal, or standard, completion time of a task for analysis.

Analysis of temporal activity associated with any task execution requires consideration of many factors including the work pace and possible synergistic influence of actions on one another (Bedny, Karwowski, & Voskoboynikov, 2015). The widely accepted standard of temporal analysis in the performance of physical activities (work) is the time study. The time study is a technique of establishing an allowed time standard to perform a given task, based upon measurement of work content of the prescribed method, with due allowance for fatigue and personal and unavoidable delays. According to Meyers and Stewart (2002), the development of time standards can be defined as determining “the time required to produce a product at a work station with the three conditions: (1) a qualified, well-trained operator, (2) working at a normal pace, and (3) doing a specific task.” The earliest of research oriented time studies relied on time consuming, stopwatch-style procedures and techniques; however, since the end of World War II, there has been significant research into the development and use of basic, predetermined motion times that can be used to predict standard times for new or modified work environments (Freivalds, 2009). These predetermined motion times are the result of large sample studies of diverse operations that culminate in tabulated guidelines and instructions on their use. Examples and derivations of current, commonly used, predetermined time systems are listed in Table 6.

Table 6: Predetermined Time Systems

Systems name	System description	Relevant applications
MTM systems (Maynard, Stegemerten, & Schwab, 1948)	Methods-Time Measurement	MTM-3, MTM-UAS, MTM-MEK and MTM-B (tasks between 1 to 5 minutes and longer)
MOST (Zandin, 2003)	Maynard Operation Sequence Technique	MaxiMOST (more than several minutes), (non-repetitive operations)
MODAPTS (Heyde, 2001)	Modular Arrangements of Predetermined Time Standards	Easy to apply system for setting labor standards
EASE (DoD, 1997)	Work measurement and time standards (MIL-STD 1567A compliant)	Example of commercially available, computerized, predetermined time systems

Although there is still some uncertainty across the research community today as to the validity of aggregating discrete, predetermined times to identify cumulative task times, predetermined time systems have become widely used across a broad range of industries (Abbott, Liu, Chua, & Chang, 2011; Yadav, 2013; Sabaric, Brnada, & Kovacevic, 2013). In addition, predetermined time systems have become the foundational standard for research improvements and adaptations (Christmansson, et al., 2000; Nakayama, Nakayama, & Nakayama, 2001; Kuo & Wang, 2009; Kuhlang, Edtmayr, & Sihn, 2011) as well as an accepted standard for comparison (Razmi & Shakhs-Niyaee, 2008).

Once predetermined time systems, or another viable method such as empirical studies or subject matter expert (SME) input, have identified the cumulative task performance time within a complex STS, a typical normal distribution can be used to describe inter-individual performance variability as shown in Figure 6.

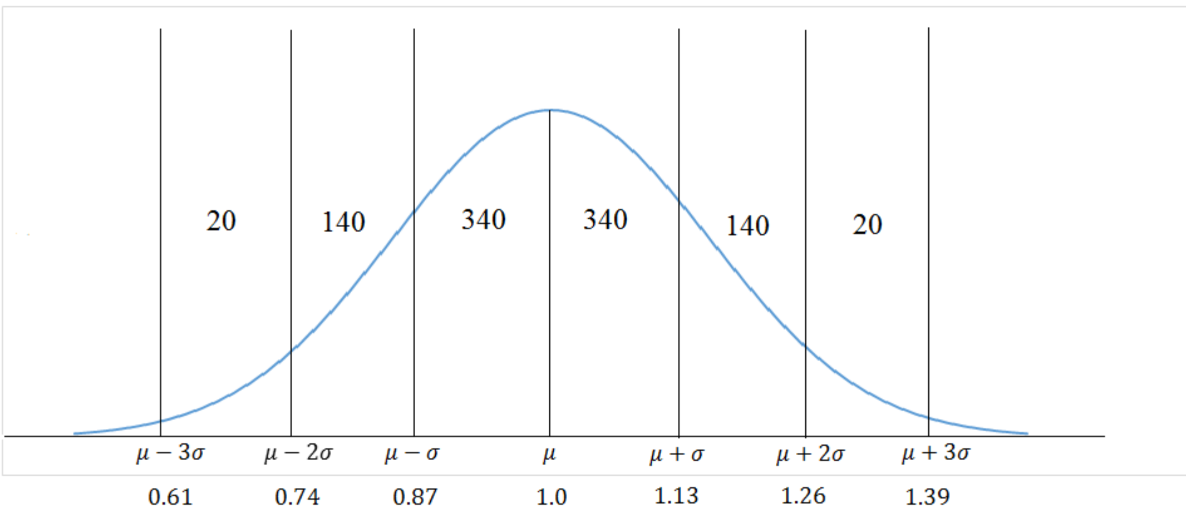


Figure 6: Expected Human Performance Time Distribution
(Adapted from Freivalds, 2009)

With inherent expectation of inter-individual variability for a random sample of 1000 employees, Figure 6 shows the frequency distribution of performance, with over 997 cases on average falling within three-sigma limits of the mean. “Based on the ratio of the two extremes (1.39/0.61), the best individual would be more than twice as fast as the slowest individual” (Frievalds, 2009, p. 440). This expectation of standard performance becomes the starting reference for considering the impacts of PIFs on the physical abilities of individuals in the completion of a task.

2.3.6 Human Temporal Variability Summary

Time variability in human performance is a dominant and recurrent concern within the research community, particularly for human-centric complex adaptive systems like nuclear reactor, military, and air traffic control operations. Dependent on the type of abilities (cognitive, psychomotor, physical) being exercised in the performance of a particular task, a variety of time distributions/equations are used to describe the response variability. This research has been evolutionary in nature and generally recognizes that human temporal performance is cumulatively and synergistically dependent on the discrete elements of individualized human abilities as well as on the factors that impact those abilities. As a result, there exists an acknowledged need for an understanding of the mechanisms describing human task performance to include a defined taxonomic structure categorizing both human abilities within a given task design and the task performance influencing factors. Substantial research has been completed in developing each of these; however, a general consensus as to the best taxonomic categorization of the bounds of abilities and factors to be considered has yet to be attained.

CHAPTER THREE: METHODOLOGY

In computer science, a methodology is a set of guidelines for covering the whole end-to-end process of agent based model development, both technically and managerially. The design methodology provides the process with guidelines and architecture that can be used to construct the system, its components, and the interactions between the components (Siebers & Aickelin, 2008). Implementing an Agent Based Model (ABM) to simulate and assess U.S. Navy shipboard operations requires modeling dynamic and coherent crew relationships representative of real shipboard performance. Once the elements of crew member capabilities, their relationships and interactions, and influences of the external shipboard environment are known, a representative model can be built and benchmarked. The representative model must be built on the foundation of a suitable architecture and employ the use of empirical or subject matter expert derived data to achieve satisfactory analytical outcomes. Once built, the representative model provides the foundation from which all facets of crew temporal variability impacts can be investigated.

3.1 ABM Architecture

Agent Based Modeling and Simulation is a relatively recent modeling method based on object oriented programming, Unified Modeling Language (UML) architectures, and the use of statecharts. It allows placing agents in an environment and monitoring their individual behavior as well as their interactions with other agents within the environment. For this research, agents were integrated into the representative model through the creation and use of workload performance statecharts representing the agents as well as their interactions with the environment and each other. Workload performance statecharts indicate what states the crew member can be in and what triggers state changes for them. Triggers can change the internal state of the crew member or establish state changes in other entities around the crew member. States can have different levels of importance and can result in a variety of outcomes including execution of timing impacts (both positive and negative) or interruption of other state changes. Trigger state knowledge incorporated within the statecharts is typically represented as formulas, rules, probabilities, or procedures. A conceptual example of a crew member statechart is shown in Figure 7.

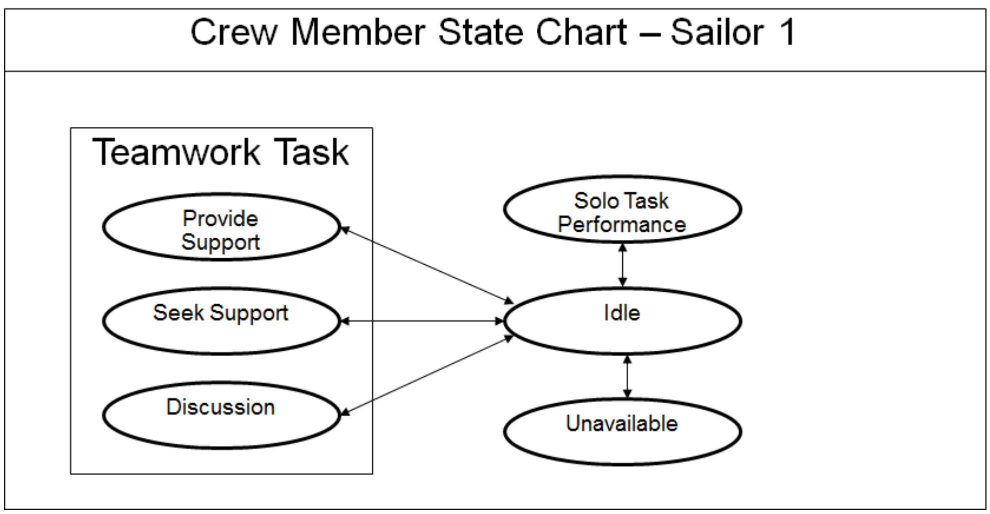


Figure 7: Conceptual Crew Member Statechart

Once the empirical data and crew member statecharts are integrated into the architectural framework for the baseline model, the model can be run to provide indications of errors or inconsistencies. Errors and inconsistencies that are uncovered will be addressed through debugging. Debugging is an iterative process aimed at uncovering and correcting errors in the model implementation (Balci, 1998). Debugging iterations are carried out in four steps: test the model to detect errors, determine the cause of the error (the bug), identify required changes or modifications, and execute the changes. The iterative process continues until no additional errors or inconsistencies are discovered. The debugging process will be performed following each incremental addition of functionality and/or complexity and will ultimately support the verification of the model.

The number and complexity of crew relationships and interactions represented by statecharts for the selected small boat defense scenario required an incremental shipboard to computer application model development strategy that added functionality through stages. For this purpose, AnyLogic ABMS software proved to be a very capable object oriented modeling and simulation tool. AnyLogic is a proprietary simulation software that employs Java language for the definition of complex structures and algorithms, combines three main simulation methodologies (system dynamics, discrete-event, and agent-based modeling), and supports different types of simulation experiments including parameter variation and optimization. The main building blocks of an AnyLogic model are active agents that have their own unique functionalities and interactions within their environment. The behavior of an agent in AnyLogic is defined using the statechart methodology described earlier.

3.2 Data Availability and Use

Once representative empirical or subject matter expert-derived data is provided, the crew member statechart architecture can be integrated into the ABMS framework for analysis. The U.S. Navy, as part of crew right-sizing efforts, established a comprehensive Human System Integration (HSI) task repository. The HSI task repository database is a comprehensive listing of crew performed tasks consisting of 78 attributes for each task. The tasks were collected and adjudicated using both empirical data and subject matter expertise. One of the key attributes for each task is idealized performance time, given in terms of a maximum time, a minimum time, and a mean time. The HSI task repository served as the primary analytic tool underlying the quantitative specification of workload in the crew design and analysis for all DDG-1000 ship building program usability tests. For this research, data mining was used on the database to identify representative times for each small boat defense watch team agent. Then, with a given time distribution for each task, AnyLogic agent based simulation was used to assess the impacts on temporal variability within the scenario as a function of agent capabilities and influencing factors.

3.3 Agent Development and Use

An overview of the shipboard small boat defense scenario, showing the agents simulated and information passed between them, is depicted in Figure 8.

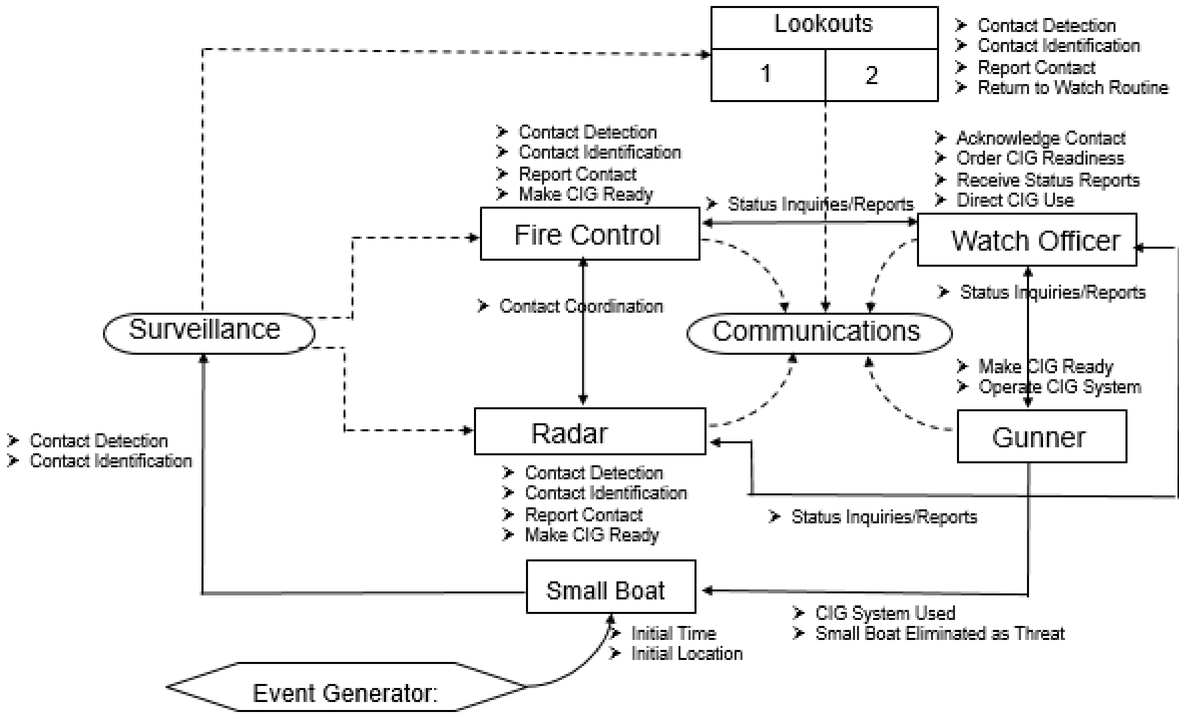


Figure 8: Agent Functionality and Information Passing for Small Boat Defense Scenario

In Figure 8, direct interactions between sailors are indicated with solid lines, whereas functionality and/or communications involving two or more sailors are indicated by dashed lines. The agents and their behavioral assumptions modeled are described further in Table 7.

Table 7: Modeled Agents and Their Behavioral Assumptions

Agent Name	Description
Watch Officer	The Watch Officer is the Commanding Officer's direct representative in supervising personnel on watch and directing them as necessary to support the safe and effective operation of the ship. For the small boat defense scenario, their responsibility is to assure that small boats, once identified, are tracked and eliminated as threats. This includes communicating with the watch team to assure that the Close-In Gun System is ready and permission to fire the Close-In Gun system is given if needed.
Lookout 1 **Lookout 2**	The Lookout maintains a continuous watch of the sea and reports any kind of hazard that can cause harm to the ship. Lookouts give their uninterrupted attention at all times to the ship's navigation and inform the Watch Officer about other ships or hazards to navigation. For the small boat defense scenario, they are key members of the watch team, along with the Radar Operator, likely to report the detection of a small craft to the rest of the watch team.
Fire Control	Fire Control is responsible for all operational aspects of the computer and control mechanisms used in weapons systems and related programs. In the small boat defense scenario, Fire Control assigns electronic target trackers to contacts upon orders from the Watch Officer to make the "Close-In Gun System Ready."
Radar Operator	The Radar Operator uses radar and other electronic equipment for the collection, processing, display, evaluation, and dissemination of small boat contact information to the Watch Officer and other members of the watch team. They are key members of the watch team, along with the lookouts, tasked with the responsibility of locating and identifying new contacts. Additionally, upon orders to make the "Close-In Gun System Ready," the Radar Operator coordinates with the other members of the watch team to assure that the target is assigned to an electronic tracker and that pertinent information is passed to the Gunner.
Gunner	Gunners are responsible for the operation of gun mounts and other ordnance equipment, as well as small arms and magazines. In the small boat defense scenario model, the Gunner receives orders from and interacts with the watch officer and other members of the watch team. The Gunner performs the final operations to eliminate the small boat as a threat. The simulation ends once the Gunner's actions are complete.

In describing agent (crew member) actions, specific roles and responsibilities across the shipboard command and control domain were considered. The actions of specific watch team agents are orchestrated by the Watch Officer in response to developing events. These interactions, following the identification of a new contact, require full representation of crew management processes, including communication and collaboration. This system results in the ability to evaluate the Watch Team and Watch Officer from a fully detailed human performance behavioral model perspective. This model assumes the temporal behavior of each individual watch team agent, as needed, and integrates it with the temporal behavior of every other watch team agent. This approach allows the mundane aspects of crew member interactions with each other as well as the simulation environment to focus on those tasks of particular concern to the scenario. As the behaviors associated with response are completed, the internal representation is stored, and the human performance model elicits temporal performance data from each watch team agent. Conceptually, this approach takes advantage of the larger simulation environment and the basic shipboard processes implicit in the dynamics of crew response to a small boat threat contact. The memory of each watch team interaction is inherent in, and provides the basis for, following actions. Pragmatically, this approach supports a detailed, expandable model of basic shipboard actions and the performance influencing factors affecting the temporal variability of watch team agents performing those actions.

3.4 Analytical Approach Development and Implementation

Workflow is the process through which a piece of work passes from initiation to completion. In the shipboard context, task workflow identifies task origination, the mechanisms by which it propagates through the ship/crew, and the manner in which it is resolved. The resultant impact of workflow is the burden of workload placed on the crew. Workload, as defined by Hart (1988), is summarized as the demand placed upon people in terms of behavioral response to events, communications, and interactions between humans and technology. Workload analysis and assessment is conducted on the assumption that an increase in task demand results in decreased human performance and a subsequent increase in task completion time. The increase in task demand can result from a wide range of factors including any of the environmental factors the sailor agents are subjected to.

Many diverse and relevant approaches to workload determination, as well as human temporal variability and performance influencing factors, are found in literature and were considered for use in development of the ABMS architecture. Some of the more relevant approaches are:

- Gregoriades and Sutcliffe (2008) provided an integral component of estimating scenario completion time using probabilistic measures of human performance with subjective estimates of workload
- Mason (2005) focused on the development of probability density functions to represent performance variability and worker activity times in task response
- Dougherty and Stutzke (1997) looked at quantifying time impacts on off-normal events using the stochastic model of time reliability correlation (TRC)
- Reer (1994) developed a time distribution as a function of time-dependent error probability

Given the objective to design and develop an ABM representing the stochastic nature of human task performance, probabilistic uncertainties were used to identify the relative likelihood of good task performance by the agent and the temporal variability impact of influencing factors. The two components used in assessing the likelihood of good operational performance were the task topology and conditional probability tables (CPT). The topology identified the qualitative part of the model with respect to various PIFs considered, and the CPT provided the quantitative causal dependencies in terms of conditional distributions.

Accurate topology development requires identification and consideration of the many factors affecting human performance. For this research, Roth’s (1992) taxonomic structure, as presented earlier, was used to partition these factors into categories with common characteristics and properties. These categories, once developed, were used to model aspects of interest in the assessment of human operational performance. The topological architecture developed for this research used individual sailors, functioning as part of a watch team, as the agents of interest. Their performance capabilities serve as one of the determinant factors for the representative shipboard scenario.

The other factor to be considered is the task support influencing factors. Task support influencing factors, in the context of this research, consist of organizational factors affecting situational awareness and the manner and method of task completion. Examples of task support influencing factors are noise levels, training, illumination, instrument configurations and displays, team structure, machine controls, communication, and ease/distance of required movement.

Once crew capabilities and task support factors have been defined by existing empirical data and/or subject matter expertise, the topology as shown in Figure 9 can be used to construct the representative ABM.

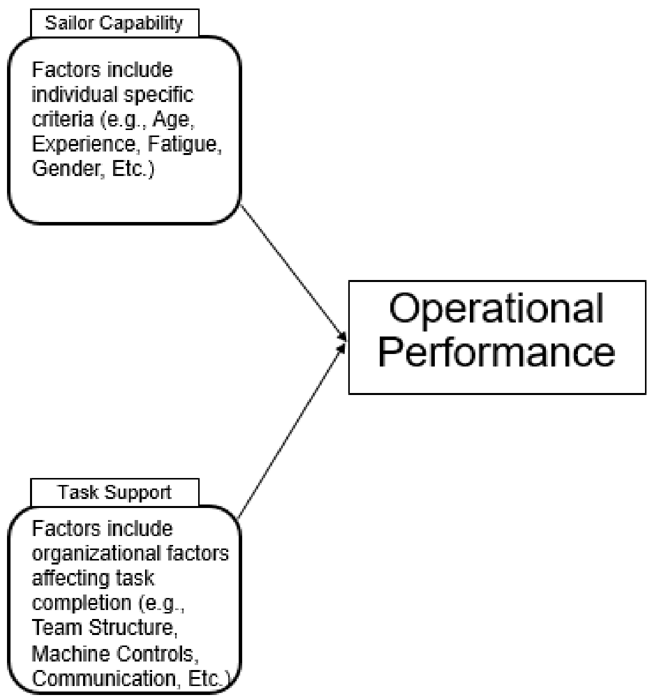


Figure 9: Topology Model for Use in Agent Based Model Development

As mentioned earlier, CPT development enables assessment of the likelihood for good operational performance in a given task and relies on knowledge of the relevant causal dependencies in terms of conditional distributions. For this research, watch team capability and task support are the causal dependencies of concern. Typical CPT development requires a simultaneous and exhaustive pair-wise combination analysis for all causal dependency likelihoods. This approach becomes intractable for a large number of considered factors.

Although this research considered only two causal dependencies, adhering to the goals laid out by the research questions requires development of a generalized model capable of considering a significantly larger number of causal dependencies. Therefore in this research, the noisy-OR approach, as defined by Pearl (1988), is used to consider conditional probabilities for each causal dependency in turn, rather than considering each of them simultaneously. This is possible because in noisy-OR gates, each causal factor is independent of any other causal factor (Lemmer & Gossink, 2004; Onisko, Druzdzel, & Wasyluk, 2001; Onisko, Druzdzel, & Wasyluk, 2000). Thus, to represent a noisy-OR CPT, only the inhibition probability for each causal node is needed, and acceptable cumulative impact probabilities can be achieved by specifying the impact each cause individually has on performance. Noisy-OR techniques have been shown to provide reasonable outcomes relative to the full implementation of exhaustive pair-wise comparison and provide a more tractable solution that is reasonably incorporated into equations built to support ABMs (Druzdzel & van der Gaag, 2000; Friedman Nir: Goldszmidt, 1996; Zagorecki & Druzdzel, 2004).

The term *noisy*, in this approach, indicates that causal interaction is not deterministic, in the sense that any cause may produce the effect with some probability, but the presence of a particular cause does not guarantee the occurrence of the effect. Thus, if $x_1; x_2; x_3 \ldots; x_n$ are causes to an effect y then each of the causes has a probability (p_i) capable of producing the effect in the absence of all other causes. These constitute the impact probabilities that allow the parameterization of the entire CPT using n causal dependencies where $p_1; p_2; p_3 \ldots; p_n$ represent the impact probability of the effect occurring given x_i is present and all others absent. In other words,

$$\boldsymbol{p_i = P(y|-x_1, -x_2, -x_3 \ldots x_i \ldots - x_n)} \tag{3}$$

and the probability of y given a subset x_j of x_is is calculated using:

$$\boldsymbol{P(y|x_j) = 1 - \prod_{i:x_i \ni x_j}(1 - p_i)} \tag{4}$$

(Gregoriades & Sutcliffe, 2008)

Once the impact probabilities (p_i) have been decided, the noisy OR method is used for generation of the CPT. Impact probabilities for a wide array of PIFs, including those of interest for this research, have been developed by the Human Reliability Analysis (HRA) research community. For this research, impact probabilities for each PIF were based on the THERP database from the handbook of human reliability analysis (Swain & Guttman, 1983). Figure 10 depicts the developed architecture that supported ABMS development using impact probabilities for crew capability and task support as identified in the cited literature.

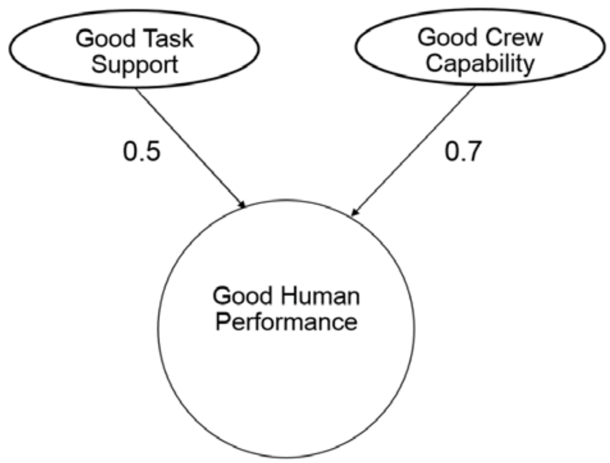


Figure 10: Impact Probabilities Relationships to Support ABM Development (adapted from Gregoriades and Sutcliffe (2008))

Conditional Probability Table development was accomplished assigning one of two possible conditional states (high or low) to the variables of crew capability and crew task support. As shown in Figure 10, impact probabilities for each of these conditional states result in inhibition probabilities of 0.3 and 0.5 respectively for crew capability and task support. Therefore, the probability that sailor performance is high, given that both task support and sailor capability is high, is given by:

$$\boldsymbol{P_{Good}}(\text{Good task support and Good Sailor Capability} = \text{true}) = \mathbf{1} - [(\mathbf{1} - \mathbf{0.5})(\mathbf{1} - \mathbf{0.7})] = \mathbf{0.85} \quad (5)$$

This in turn implies that the probability that sailor performance is bad is given by:

$$\boldsymbol{P_{Bad}} = \mathbf{1} - \boldsymbol{P_{Good}} = \mathbf{1} - \mathbf{0.85} = \mathbf{0.15} \quad (6)$$

Table 8 depicts the noisy-OR analytical development of the CPT for Figure 10.

Table 8: Application of Noisy-OR Approach for CPT Parameterization

Good Task Support	**Good Sailor Capability**	**Good Sailor Performance**	**Bad Sailor Performance**
True	True	$1 - 0.15 = 0.85$	$0.15 = (0.5)\,(0.3)$
True	False	$1 - 0.5 = 0.5$	0.5
False	True	$1 - 0.3 = 0.7$	0.3
False	False	$1 - 1 = 0$	1

(adapted from Gregoriades and Sutcliffe (2008))

The determined probability (P) of high or low operational performance by the watch team performing the task can then be used to determine task timeline and temporal variability impacts by using the following formula:

$$\text{Adjusted Time } \left(T_{adj}\right) = (P_{low}(\text{Worst Time})) + \left(P_{high}(\text{Best Time})\right) \tag{7}$$

(Gregoriades & Sutcliffe, 2008)

Best and worst completion times for each task were identified through a variety of empirical and analytical methods dependent on the given scenario as discussed in section 3.2.

CHAPTER FOUR: SMALL BOAT DEFENSE SCENARIO

Once confidence was established in the adequacy and accuracy of the developed ABM, it was used to simulate routine and non-routine shipboard scenarios to analyze crew watch team temporal variability in the performance of their duties. The small boat defense scenario was selected as a representative case for analysis based on the diversity of tasks completed, the availability of data on these specific operations, and the applicability to address a relevant U.S. Navy safety issue. The vulnerability of U.S. warships conducting littoral operations has long been a concern for the United States Navy. Rules of Engagement restrictions, as well as requirements imposed by innocent transit passage, allow potential adversaries unique opportunities to test both the engagement criteria and capabilities of U.S. Navy vessels.

4.1 Motivation

In littoral waters where United States Navy vessels routinely conduct operations, traffic density is often high, with many ferries, fishing boats, and large cargo ships maneuvering in a small area. With a host of stationary and randomly moving boats, determining a hostile action in a timely manner is difficult at best. These conditions make the identification of and defense against a hostile small craft extremely difficult (Tiwari, 2008). For these situations, the number one enemy to a Commander attempting to protect their ship against small boat attack is time. Time is central to the problem because many factors, including human response time variability, impact and compress the time needed for action. In the small boat defense scenario, the sailors are the agents of interest, and their response time variability, both on an intra- and inter-individual basis, serves as one of the performance measures considered for this research. The other performance measure considered in this research is the task support influencing factors. Examples of task support influencing factors are noise levels, training, illumination, instrument configurations and displays, team structure, machine controls, communication, and ease/distance of required movement. The elements of crew member capabilities, their relationships and interactions, and influences of the shipboard task environment provide the architectural framework for crafting a representative agent based model.

4.2 Scenario Design

This scenario investigates the impacts of emergent behavior and workload performance time variability associated with a U.S. Navy vessel watch team during small boat defense operations. In these operations, the watch team task begins with the sighting of a small boat and direction from the Watch Officer (WO) to verify readiness of the Close-In Gun (CIG) system. The awareness of a new contact and corresponding direction from the WO propagates through the ship via notification to the watch team, and is complete once the small boat is mitigated as a threat as shown in Figure 11 (critical path shown in yellow).

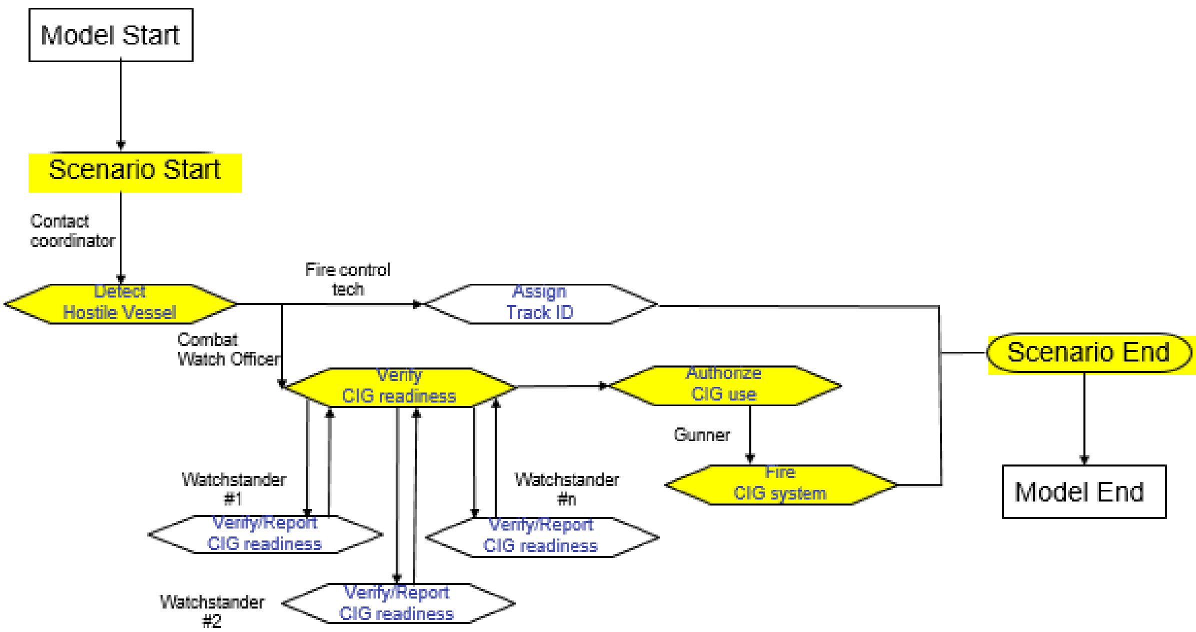


Figure 11: Small Boat Defense Scenario

A typical area that could be used for the model is the Mediterranean Ocean. This location clearly supports visualization of routings, chokepoints, and traffic densities critical to analysis of force protection in littoral waters. Figure 12 shows a representative map of the Mediterranean indicating typical traffic densities and outlining the areas of navigable water.

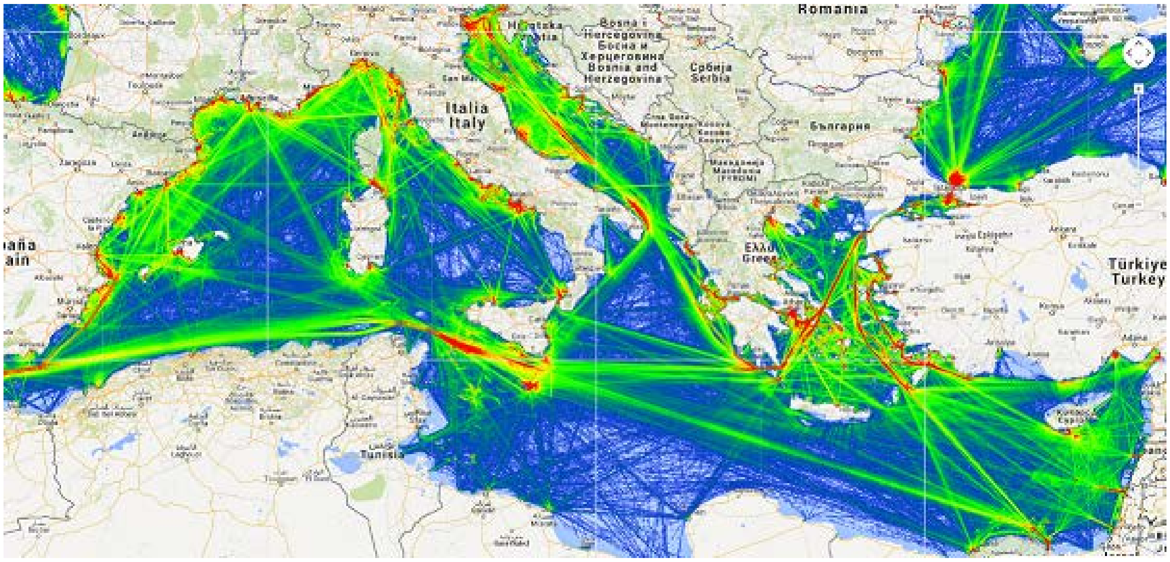


Figure 12: Mediterranean Ocean Ship Routings, Chokepoints, and Traffic Densities

The small craft of interest for this scenario is randomly created with a route-generic distribution to support assessment of a task timeline and impacts of crew temporal variability on sailor workload performance. For this scenario, a detection rate of 100 percent and a false alarm rate of zero percent were assumed. In addition, the environmental conditions for this scenario assumed unlimited visibility, low sea state (0-3 ft.), and adequate water depth to support maneuverability. When a small craft is detected by a member of the watch team, a consistently reliable alert is communicated to the crew to initiate action from a perceived “threat contact” approach and ultimately results in the Gunner firing on the small craft. Although there is continual vigilance and alerting, for purposes of the ABMS effort, only the first alert issued is acted upon by the crew. Future versions of the model can be adapted to elaborate on crew responses to a variety of detection rates, multiple detections, evaluation of formation steaming effects, and path planning in a dense environment.

4.3 Crew Behavior

Crew behavior, for the small boat defense scenario, follows standard operational procedures for littoral water navigation. The model of crew behavior uses scripts as its normative standard, generated based upon the experience of the author, a twenty year veteran in U.S. Navy operations, and the tailored use of a comprehensive Human System Integration (HSI) task database developed by the U.S. Navy. While many factors impact crew temporal response, only inherent sailor capability and the impact of task support were considered as performance influencing factors for this analysis. It was found that crew response varies with each of these factors. The details of the shipboard operations to include directive verbal orders and subsequent crew response were considered in full. The model defines a variety of crew responses to these catalysts of action in the context of different operational tasks: identifying and reporting new contacts, readying the CIG System, and firing the CIG System.

4.4 Small Craft Behavior

As described by Tiwari (2008), small craft tactics can typically be broken into three distinct modes: Attack, Distraction/Diversion, and Surveillance. The first threat behavior is a direct attack. In this mode, threat craft move from the point of origin directly toward their target and attack as soon as they enter the attack range. The second behavior is implemented by using one or more threat craft to distract/divert their target’s attention, through harassment, and then launch an attack from another craft. In the surveillance mode, a threat craft loiters in the vicinity of the target and switches tactics to attack upon order from a coordinating authority. In this scenario the direct attack was the only behavior modeled; however, all of the behaviors are representative of possible real world scenarios and could be implemented into the ABMS process.

4.5 Weapons

Weapons effects are not modeled for the purposes of this simulation. However, the framework for the inclusion of real world weapons parameters is included, allowing for future work to model and analyze these effects with the provision of weapon system-specific data.

4.6 Radar

Although hostile action and/or hostile intent cannot be discerned from a radar picture in this environment, radar identification of a potential small craft threat is likely one of the first watch team queuing mechanisms. Both surface search and air search radars were assumed operational in the modeled scenario and capable of providing the standard measures for threat assessment such as speed, angle of approach, and location for up to 45 contacts. The Watch Team, once alerted of a fast-moving inbound contact, initiates an assessment and engagement sequence to determine the intent of the small craft. However, high speed alone is not an unmitigated qualifier as to intent. Vessels may legally operate at fast speed while loitering contacts offer little to no information as to intent and can launch a very effective attack from a very close range. For these reasons, the watch team has a strong reliance on visual sensors and response to radio communications between the Navy vessel and perceived threat small craft.

CHAPTER FIVE: RESULTS AND DISCUSSION

Three general categories of measures were made from this ABMS approach to U.S. Navy Small Boat defensive operations:

1. Measures of emergent systems behavior and performance. This measure includes individual task times as well as the time for the total system of tasks (lookouts, radar operator, fire control operator, watch officer, and gunner) to detect, respond to, communicate about, and mitigate the small boat threat. Workload utilization for each member of the watch team is also considered.
2. Measures of individual sailor capability and demands placed on sailors by the environment. This measure identifies individual sailor performance temporal impacts due to scenario uncertainties to include watch team task loading, communication/behavior interfaces, and restoration to a normal watch team routine.
3. Scenario impacts based on uncertainties associated with crew capability and the shipboard environment. This measure compares and contrasts the idealized model of sailor temporal performance with the model developed to consider uncertainties due to the influencing factors of sailor capability and task support.

The first category of measure deals with illustrating the mechanisms and outcomes of the developed baseline approach including the verification of model performance. The second category of measure presents the results of executing the ABMS methodology with different values for the parameters, sailor capability and task support, thus providing a study of its overall behavior from a realistic point of view. The third category of measure considers the uncertainties of sailor operational performance based on technology interfaces, inherent characteristics, and the environment, and compares them to the idealized times obtained from the HSI database.

5.1 Demonstration and Verification of the Baseline Model

The baseline model for this project was built, debugged, and verified using timeout state transitions composed of either constants and/or triangular distributions reflecting the idealized times from the HSI database. Representative baseline model development was completed to assure that "real-world" shipboard element architectures were integrated into the scenario simulation. Fully understanding the architectural framework under which each sailor will operate requires scenario specific knowledge, and awareness of workflow, decision points, and crew relationships. Scenario specific knowledge is driven by fully understanding the required functionalities of each crew member of interest (agent) for a specific performance task (e.g., detect/report small boat, verify CIG readiness, etc.). To illustrate these functionalities on a system-wide basis, a roll-up AnyLogic statechart application for the small boat defense critical path is provided in Figure 13 to indicate the required watch team actions from scenario start to scenario finish.

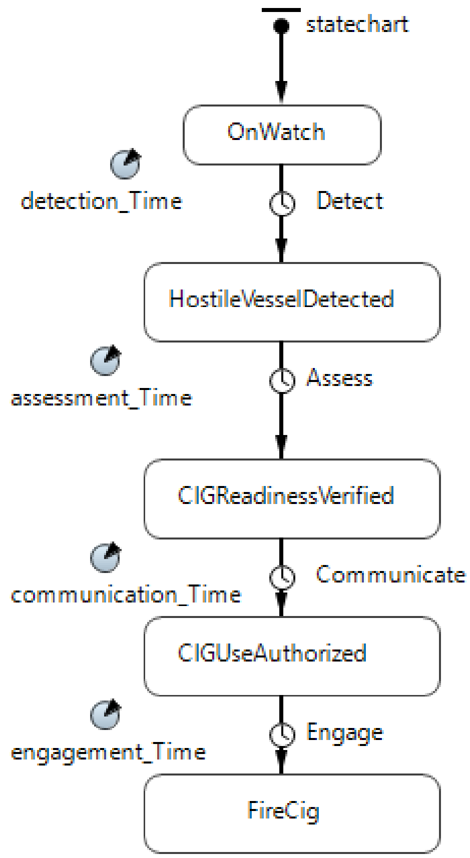


Figure 13: Small Boat Defense AnyLogic Statechart

In the detailed AnyLogic baseline model used for analyzing watch team performance, individual watch team agent actions for each state were completed and compiled into a comprehensive scenario response. In this response, the small boat contact appears on the horizon and is identified by one of four watch team agents: lookout 1, lookout 2, radar, or fire control. Once identified by the watch team agent as a potential threat contact, the small boat is then reported to the Watch Officer. Upon successful completion of the identification and reporting performance tasks, the lookouts return to their on watch function and the Fire Control and Radar watch team agents perform additional actions as directed by the Watch Officer as shown in Figure 14.

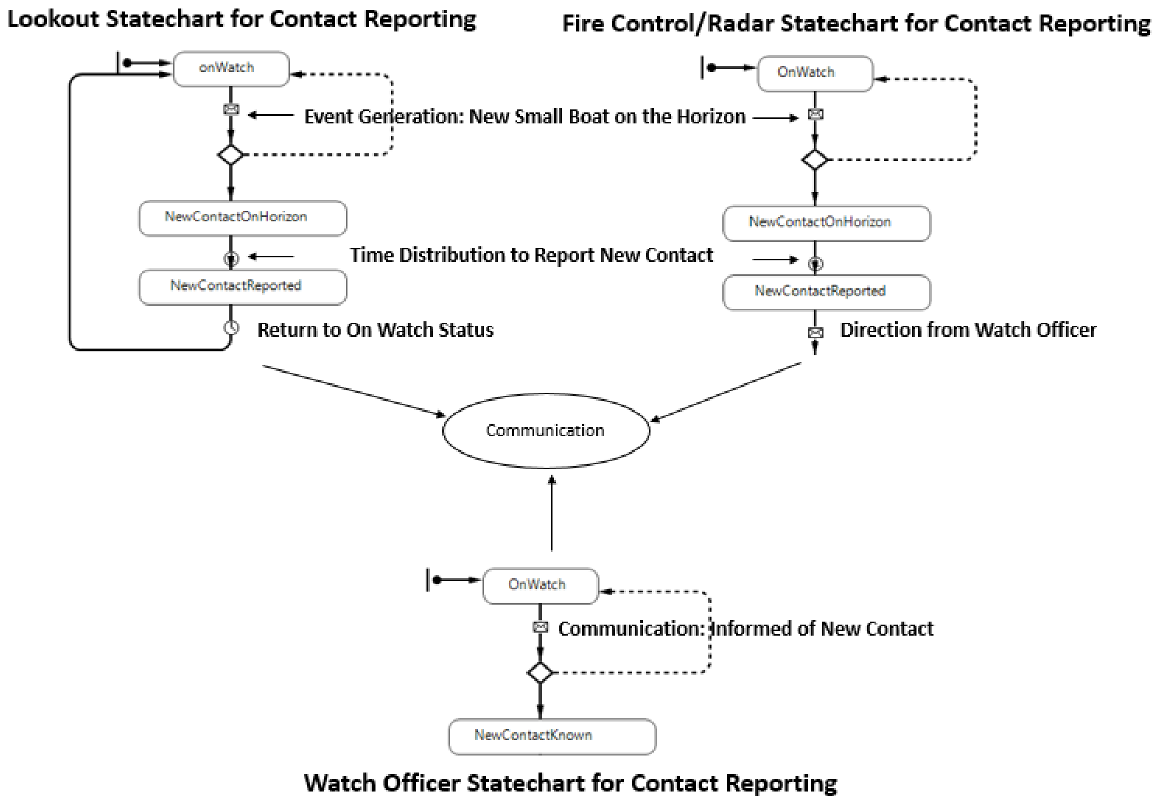


Figure 14: Individual Watch Team Agent Statecharts for Small Boat Detection

Once the small boat contact is identified to the Watch Officer, its presence is formally communicated by the Watch Officer to the rest of the watch team. The Watch Officer then directs Fire Control, Radar, and Gunner actions to verify and report the readiness of the CIG system. The performance task is not complete until all three watch team agents have successfully communicated completion of their actions to the Watch Officer as shown in Figure 15.

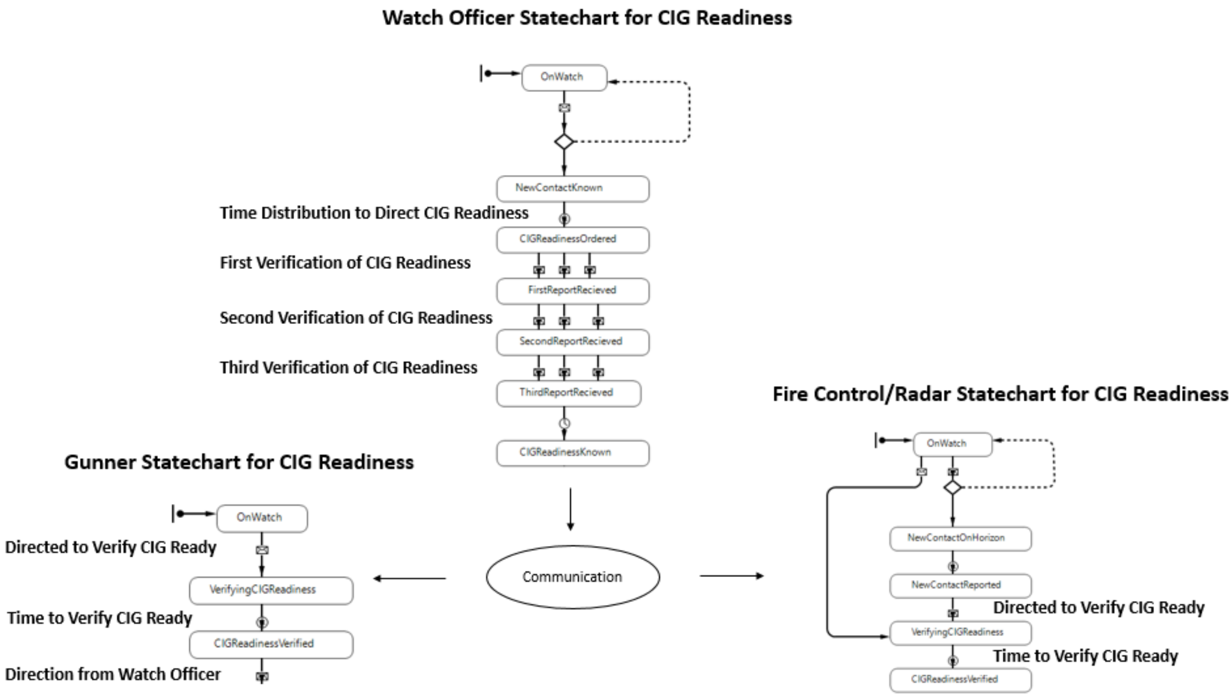


Figure 15: Individual Watch Team Agent Statecharts for Verifying CIG Readiness

Figure 15 also shows that the model provides the option for the Watch Officer to direct the actions of Fire Control and Radar via message communication, whether or not either of these two watch team agents was the initial individual to identify and report the contact. This is an essential element of the model as both the Fire Control and Radar Watch team agents will have required CIG readiness preparation actions regardless of whether they initially identified the contact or not. Upon successful completion of the CIG readiness performance tasks, Fire Control and Radar continue to perform their on watch functions, and the Gunner performs additional actions as directed by the Watch Officer. Once CIG readiness is complete and identified to the Watch Officer, the Watch Officer has the responsibility and authority to determine the necessity for CIG use and to communicate this direction to the Gunner. Once the Watch Officer authorizes CIG system use, the Gunner then uses their experience along with input from other watch team agents to initiate firing and subsequent elimination of the small boat as a threat as shown in Figure 16.

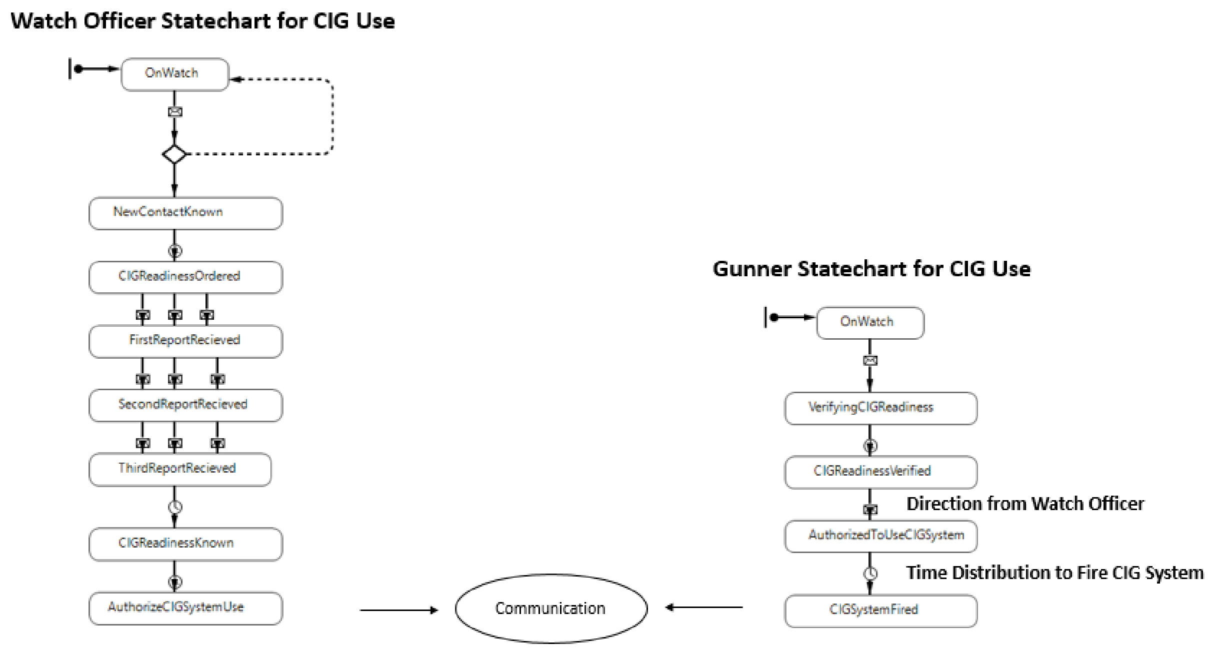


Figure 16: Individual Watch Team Agent Statecharts for CIG Use

Verification of the AnyLogic simulation estimate of sailor performance was used to demonstrate the predictive capability of the linked model, confirm the model's internal logical consistency, and verify that the model operates within reasonable expectations. Assumptions were also made that each watch team agent was capable of performing their duties and that no false alarms or missed detections occurred. For the sake of brevity, results for only a representative portion of the verification tests will be presented in this research. The following scenario demonstrates one of the 64 baseline scenarios for which verification was completed:

- The watch team is in place and ready to perform their duties 11.5 minutes after event start.
- Small boat threat contact appears on the horizon 12 minutes after event start
- Detection of small boat is at a range of 6 miles with sustained 9 knots closure – results in a critical time (T_c) of 40 minutes for elimination of the small boat as a threat
- Watch team response times used the idealized triangular distributions of Table 4
- 1000 runs for Monte Carlo statistics

Individualized and cumulative watch team tasks temporal performance as well as individualized workload statistics were the focus of this section:

- The average time to detect the small boat and communicate notification to the watch team.
- The average time for the Gunner, Fire Control, and Radar to verify CIG readiness, including communication of orders.
- The average time for the Watch Officer to authorize use of the CIG system.
- The average time for the Gunner to fire the CIG system.
- The cumulative time for the watch team to verify CIG readiness and eliminate the small boat as a threat.
- The workload utilization for each member of the watch team.

Idealized best and worst completion times for each watch team action in the small boat defense scenario were identified through data mining of the HSI database described in section 3.2 and were used in the AnyLogic model. These times are provided in Table 9.

Table 9: U.S. Navy HSI Database Task Completion Times for Small Boat Defense

Task Name	**Task Type**	**Minimum Time (mins)**	**Mean Time (mins)**	**Maximum Time (mins)**
Detect hostile vessel	Detect	5	10	15
Watch Officer directs: “Make the CIG system ready”	Communicate	1	1.5	2
Verify CIG system ready	Assess	8	13.5	19
Authorize use of the CIG system	Communicate	4	6	8
Fire CIG system	Engage	5	10	15
Elimination of small boat threat (cumulative time)	Compilation	23	41	59

The resulting verification runs of individual sailor performance for the baseline model are shown in Figure 17 and clearly indicate that the times for each task fell within the idealized time distributions given in Table 9.

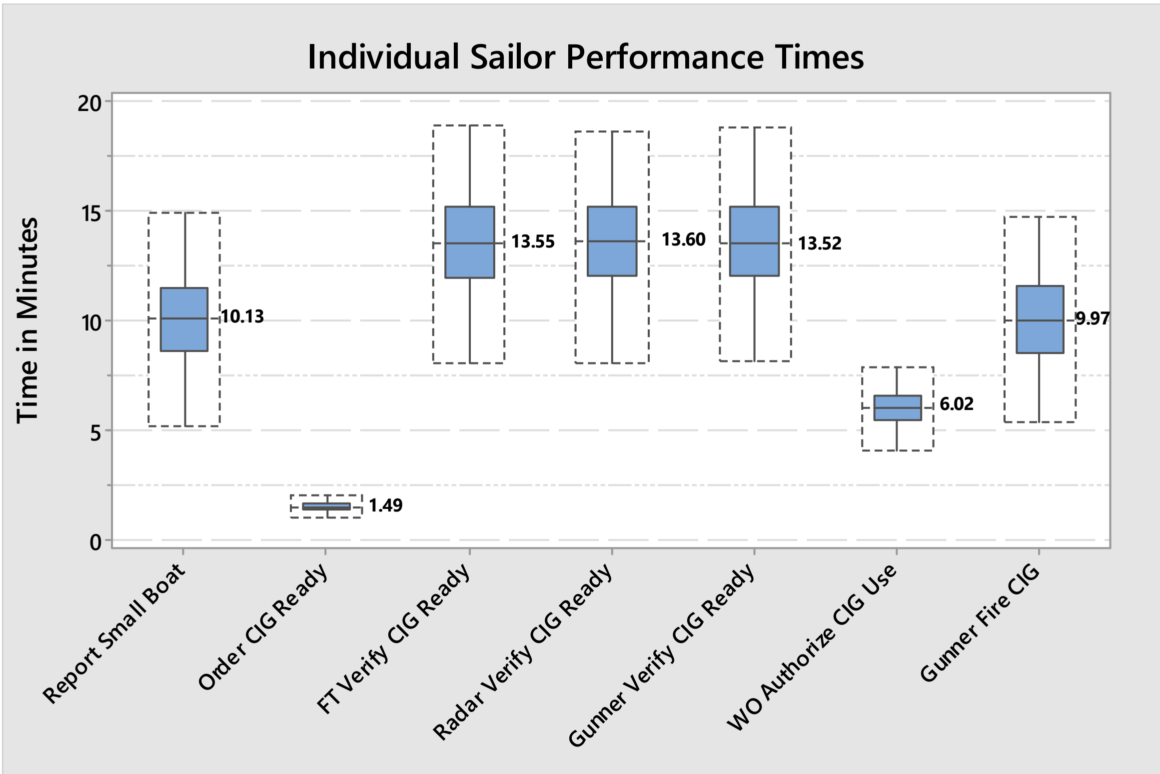


Figure 17: Individual Sailor Performance Times for the Baseline Model

Following completion of the baseline model verification for individual sailor performance, an analysis of cumulative sailor performance time for the verification of CIG readiness and overall scenario response was conducted to determine median, best, and worst completion times. Boxplots of these respective times are provided in Figure 18. The combined CIG readiness time is biased to the higher end of the band (19 minutes) as would be expected given that three stochastically determined individual reports of CIG readiness are required prior to completion of this performance task. Based on the results, watch team performance met the critical time of 40 minutes only 194 times. This equates overall to a 19.4% success rate in eliminating the small boat prior to its becoming an immediate threat to the safety of the ship and crew.

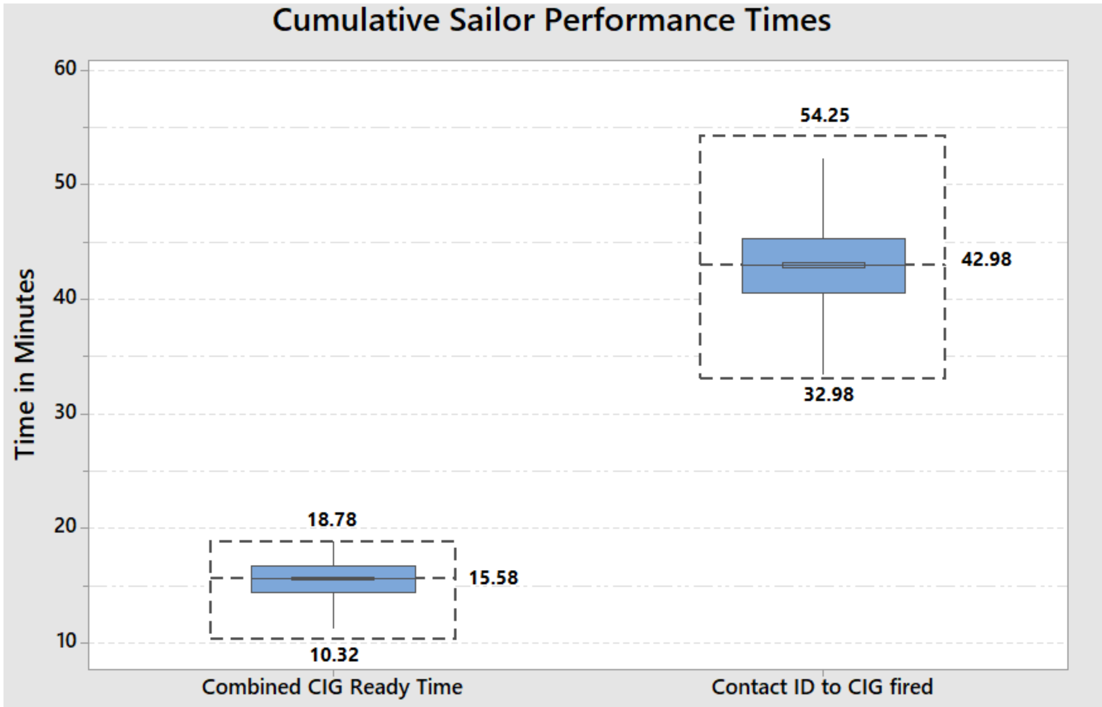


Figure 18: Cumulative Watch Team Performance Times for the Baseline Model

A histogram of the Contact ID to CIG fired time is provided in Figure 19 to provide a basis of comparison between the overall task performance timeline using the idealized subject matter expert-derived triangular distributions and the task performance timelines considering the impacts of sailor capability and task support that will be developed in the next section.

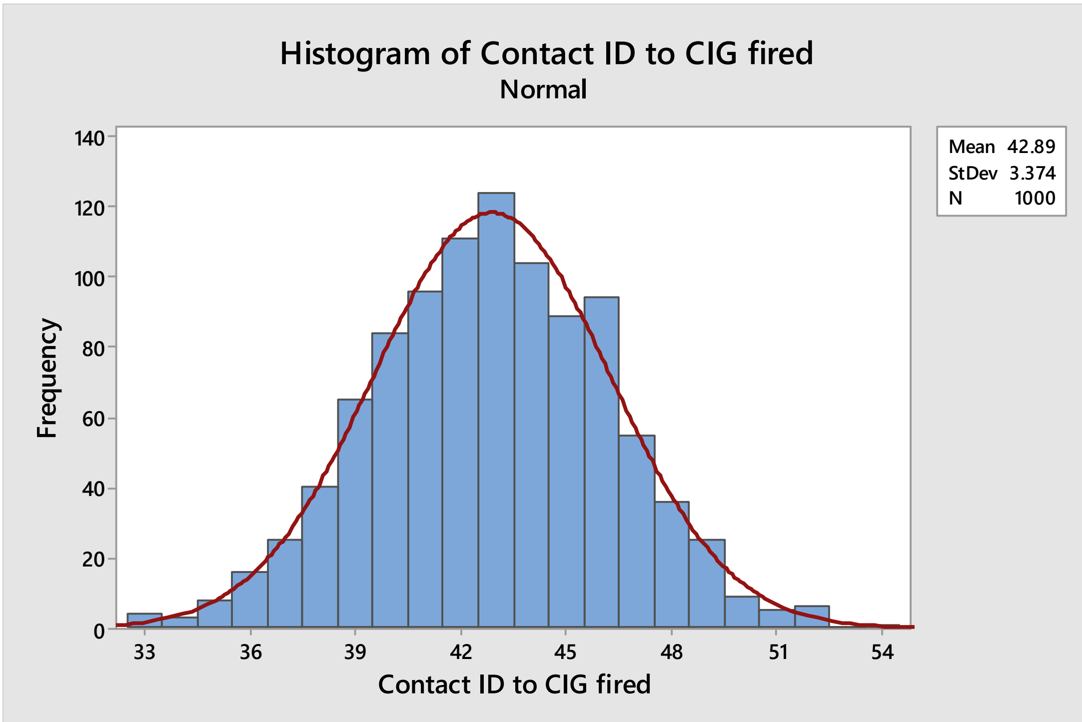


Figure 19: Histogram of Contact ID to CIG Fired Times for the Baseline Model

Workload analysis for each of the individual watch team agents was completed to identify a means of determining time sufficiency within the scenario for each watch team agent to perform their individual tasks. For the given scenario, the average workload for each watch team agent with resulting utilization is provided in Table 10. Only the watch team agent identifying and reporting the small boat is considered to have completed workload supporting this phase of the scenario and accounts for the variation in workload for the Lookouts, Fire Control, and Radar. These results are empirically consistent with observations from shipboard operations and data provided in Table 9.

Table 10: Workload Utilization of Individual Sailors

Watch Team Agent	Average Scenario Workload in Minutes	Workload Utilization
Lookout 1, Lookout 2	0 - 9.96	0 – 23.22 %
Fire Control	13.53 – 23.49	31.55 – 54.77 %
Radar	13.55 – 23.51	31.59 – 54.81 %
Gunner	25.48	59.41 %
Watch Officer	7.49	17.46 %

Baseline model results indicate that an expandable block model architecture identifying each performance task completion timeline can be successfully employed and repeated to be representative of scenarios composed of 1 through *n* tasks. In addition, for scenarios with multiple small boats and/or including parallel activities by watch team agents, workload times can be summed to clearly identify any scenarios for which the individual member does not have adequate time to complete all of their tasks as defined by workload utilizations in excess of 100%. The ability to assess a large, complex scenario, then, requires only understanding the workflow, decision points, and crew relationships for each performance task, the associated tasks' minimum and maximum performance times, and the probabilistic temporal impacts of considered influencing factors.

5.2 Crew Response Incorporating Sailor Capability and Task Support

The analysis of temporal impacts resulting from sailor task performance variability used the baseline model as the architectural foundation and incorporated variation in sailor capability and task support for each performance task. Using the architectural framework as defined by Figures 13-16, CPT development was accomplished by assigning one of two possible conditional states (high or low) to the variables of crew capability and task support as shown in Figure 10. Each individual watch team agent was assigned a likelihood of high capability and high task support, based on assumed shipboard profiles for crew rotation, crew qualification, and experience. Watch Officers were assigned a likelihood of high capability 50% of the time. This is based on the fact that most Officers onboard U.S. Navy vessels have spent considerable time onboard prior to qualifying as Watch Officer and tend to have developed specialized experience en route to the Watch Officer position. Lookouts, Gunners, Radar, and Fire Control were assigned a likelihood of high capability 30% of the time. This assignment of likelihood is based on the fact that watch team sections are typically weighted towards junior, less experienced personnel and can be composed of non-specialized personnel, particularly in the case of the lookouts. Task support was nominally assumed to be 50% for each of the watch team agents based on the restrictive task performance locales onboard ship and environmental factors, particularly given the relative high task level index of difficulty. Likelihood assignments for both sailor capability and task support can be easily varied within the developed model. With known likelihoods for sailor capability and task support, impact probabilities from the THERP database were used, as discussed in Section 3.4, to calculate adjusted performance times (T_{adj}) using the best and worst case completion times from Table 7 and Equation (7). These adjusted times were then used to define the time distributions associated with each sailor performance task required to complete the entire Small Boat Defense Scenario. Figure 20 illustrates the java script coding required to define the temporal transitions based on CPT development for the Fire Control/Radar watch team agent. This CPT development was completed for each of the watch team agents to support the analysis of temporal impacts resulting from variations in sailor operational performance.

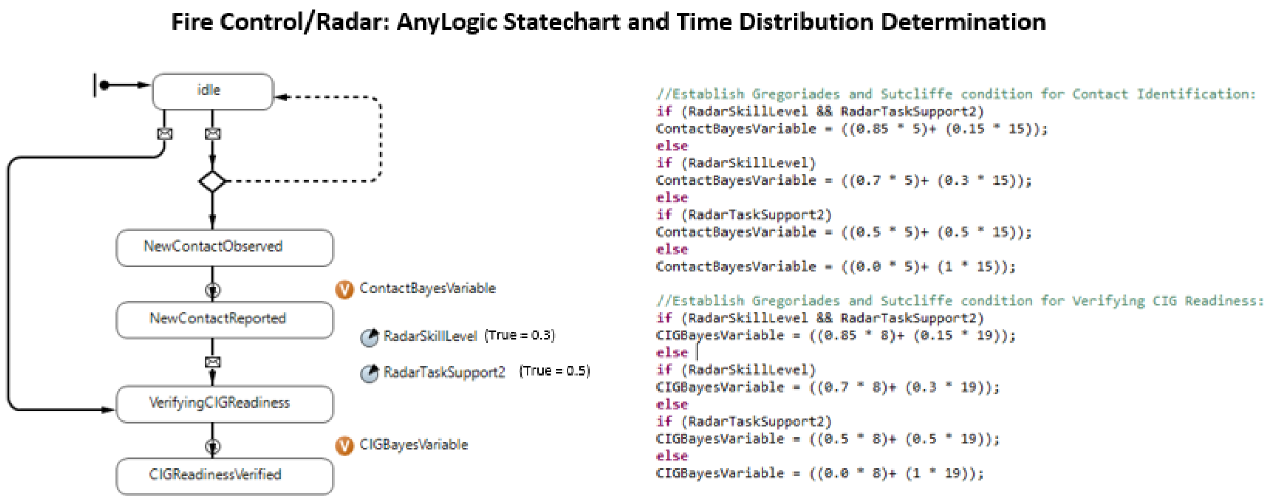


Figure 20: CPT AnyLogic Application for Crew Capability and Task Support Impacts

Full development of the analytical model enabled assessment of the temporal impacts associated with the causal dependencies of concern: watch team capability and task support. Individual temporal performance resulting from these impacts is shown in Figure 21.

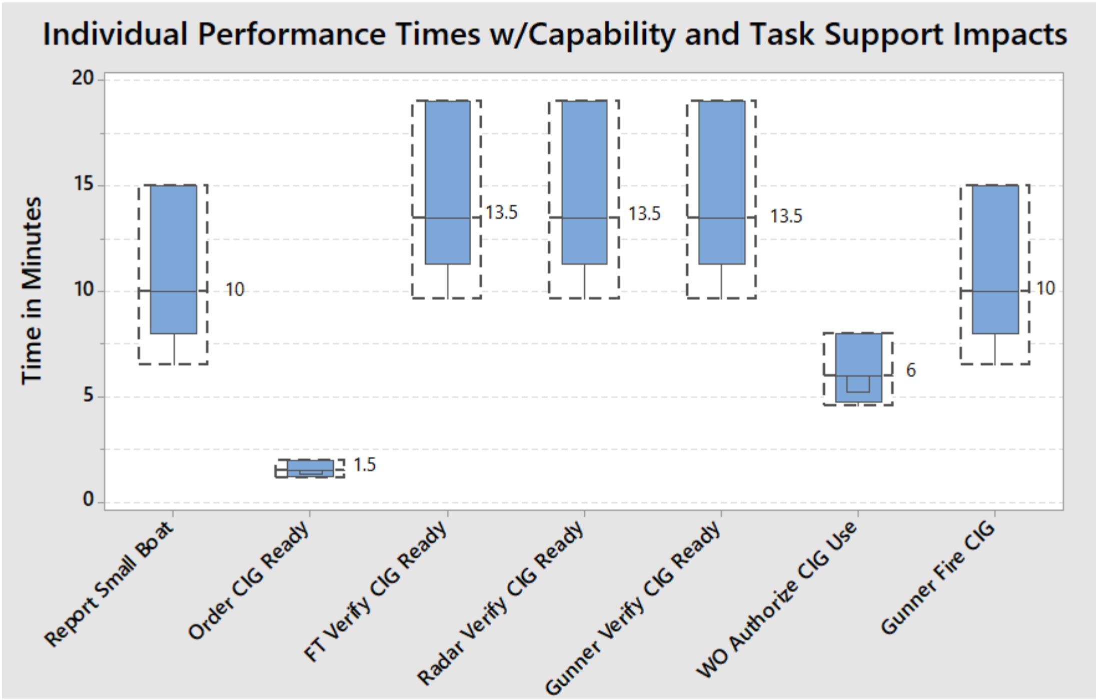


Figure 21: Individual Performance with Sailor Capability and Task Support Impacts

As can be seen from the boxplots of individual performance, sailor operational performance variability is less and task completion is weighted toward longer times in the model defined by Equation (7). This outcome is not unexpected. Equation (7) results in only four defined completion times for each individual performance task as opposed to the random triangular distribution used for the baseline case. This reduces the output variability. Also, using Equation (7), there is no likelihood that results in a completion time equal to the best completion time from Table 4, whereas for the baseline case the best completion time can be achieved. It is worth noting that the assumptions associated with the likelihood of high capability for each watch team agent affect the outcomes but are not easily distinguishable in the individual performance boxplots. For example, the Watch Officer with a 50% assumed likelihood of high capability would be expected to demonstrate times consistent with the baseline model, whereas other watch team agents, with only a 30% likelihood of high capability, would be expected to have longer nominal completion times. The boxplots of Figure 21 do not illustrate this outcome and result in the illusion of similar performance. This illusion is a result of using the median in the boxplots illustrations. With only four possible outcomes, and the assumed likelihoods, it is expected that 50% of the outcomes would fall above and 50% below the most likely value. The means for each of the individual completion times clearly show the impact of the assumed task performance likelihoods and the finite range of time options. For example, the Radar watch team agent has a mean time of completion of 13.55 minutes in the baseline case, but the analytical model employing Equation (7) results in a completion mean of 14.77 minutes. Using a two-sample t-test results in a T-Value of −9.10 and a P-Value of 0.000. Thus, we are able to reject the null hypothesis and clearly identify that the model using Equation (7) results in a longer mean time of performance task completion.

Following completion and application of the model for analysis of individual sailor performance, an analysis of cumulative sailor performance time for the verification of CIG readiness and overall scenario response was conducted to determine median, best, and worst completion times. Boxplots of these respective times are provided in Figure 22. As seen earlier in the triangular baseline model, the combined CIG readiness time is biased to the higher end of the band (19 minutes) as would be expected given that three individual watch team agents, each with a high capability likelihood of only 30%, are required to perform all of their actions and successfully report them prior to completion of this performance task. In fact, performance task temporal outcomes using Equation (7) result in the longest possible completion time in 72.7% of the model runs. Based on the cumulative scenario results, watch team performance met the critical time of 40 minutes only 177 times. This equates overall to a 17.7% success rate in eliminating the small boat prior to its becoming an immediate threat to the safety of the ship and crew. This overall lower value of success from the baseline model (19.4%) is illustrative of the bias toward longer task completion times when the performance influencing factors of sailor capability and task support are considered.

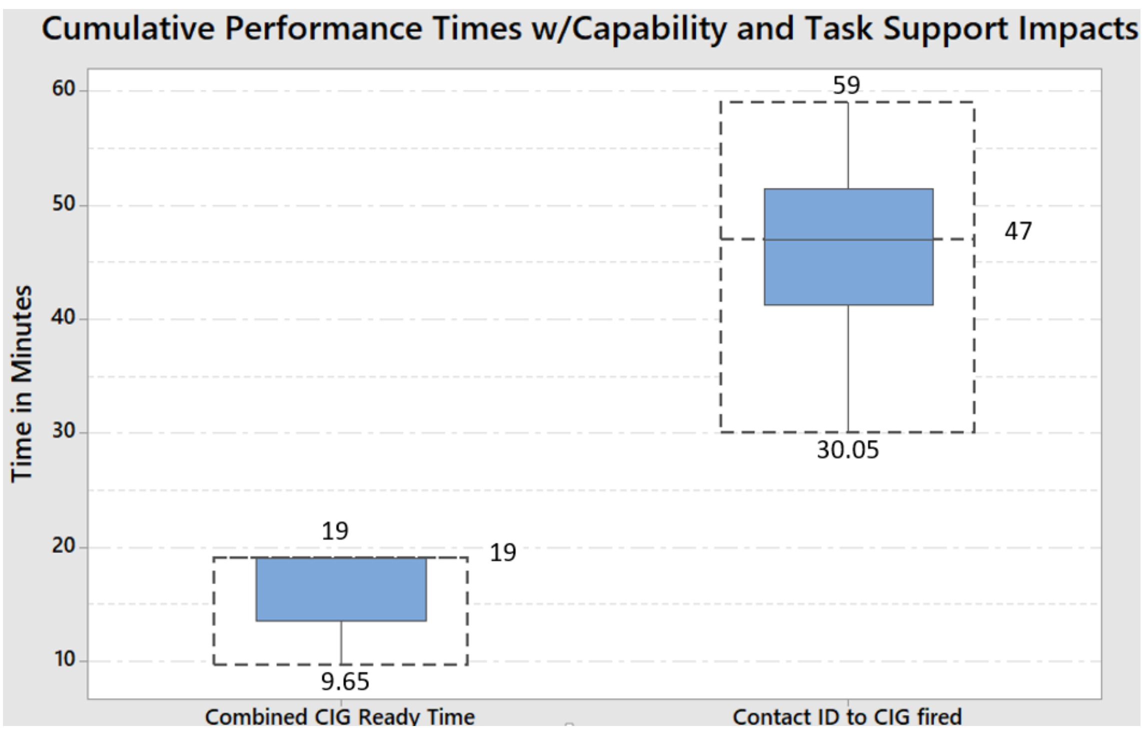


Figure 22: Cumulative Watch Team Performance Capability and Task Support Impacts

A histogram of the Contact ID to CIG fired time considering the impacts of sailor performance and task support is provided in Figure 23. It provides a basis of comparison between the overall task performance timeline using the idealized subject matter expert derived triangular distributions and the task performance timelines employing Equation (7) to identify the impacts of sailor capability and task support.

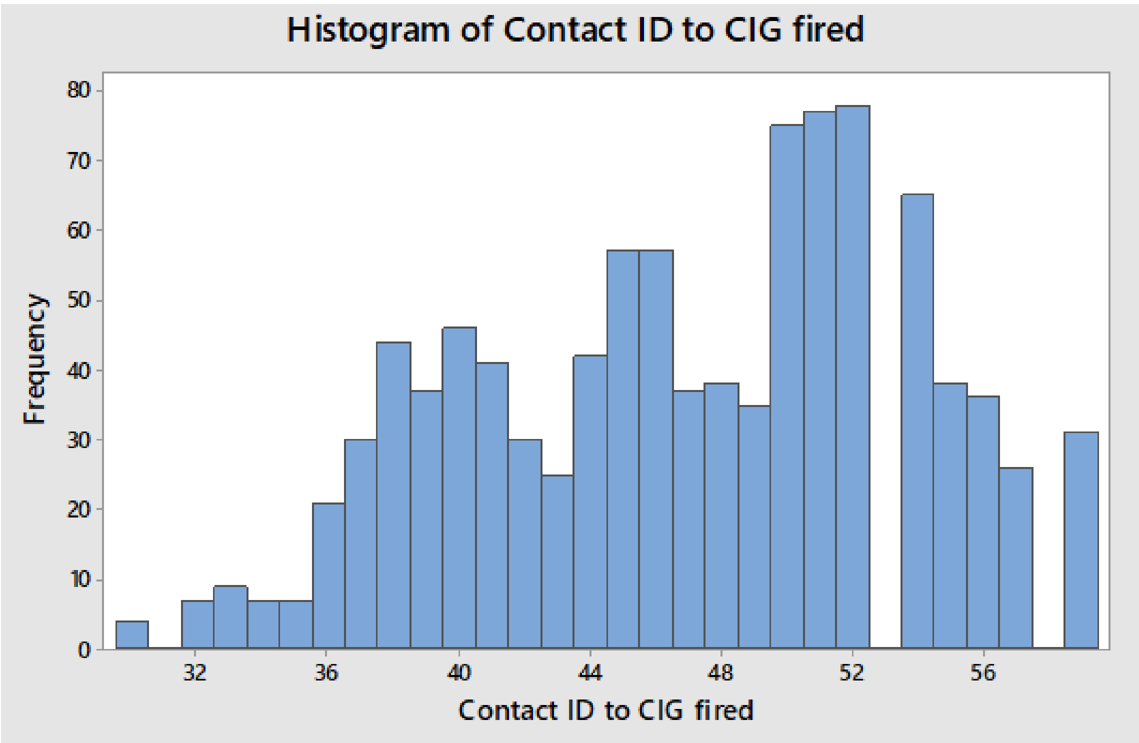


Figure 23: Histogram of Contact ID to CIG Fired Times for the Analysis Model

As shown in Figure 23, the analysis model is not nearly as uniform in output and is more heavily weighted to longer time outcomes than the baseline model. The baseline model output is more clearly representative of a normal distribution as would be expected given that the triangular distribution input was composed of a minimum, maximum, and most likely.

Similar to the effort in completing the baseline model, workload analysis for each of the individual watch team agents was conducted to identify a means of determining time sufficiency within the scenario for each watch team agent to perform their individual tasks. The average workload for each watch team agent with resulting utilization is provided in Table 11. As before, only the watch team agent identifying and reporting the small boat is considered to have completed workload supporting this phase of the scenario. This difference in assignment accounts for the variation in workload for the Lookouts, Fire Control, and Radar. These results are empirically consistent with the results from the baseline model and data provided in Table 9.

Table 11: Capability and Task Support Impacts on Workload Utilization

Watch Team Agent	**Average Scenario Workload in Minutes**	**Workload Utilization**
Lookout 1, Lookout 2	0 – 10.93	0 – 23.40 %
Fire Control	14.36 – 25.29	30.74 – 54.14 %
Radar	14.77 – 25.70	31.62 – 55.02 %
Gunner	25.31	54.18 %
Watch Officer	7.49	16.04 %

Given that the analytical model uses the baseline model architecture, it can also be successfully employed and repeated to be representative of scenarios composed of 1 through *n* tasks. Similarly, for scenarios with multiple small boats and/or including parallel activities by watch team agents, workload times can be summed to clearly identify any scenarios for which the individual member does not have adequate time to complete all of their tasks as defined by workload utilizations in excess of 100%. The ability to assess a large number of tasks requires only duplication of each input block and knowledge of the associated tasks' minimum and maximum performance times.

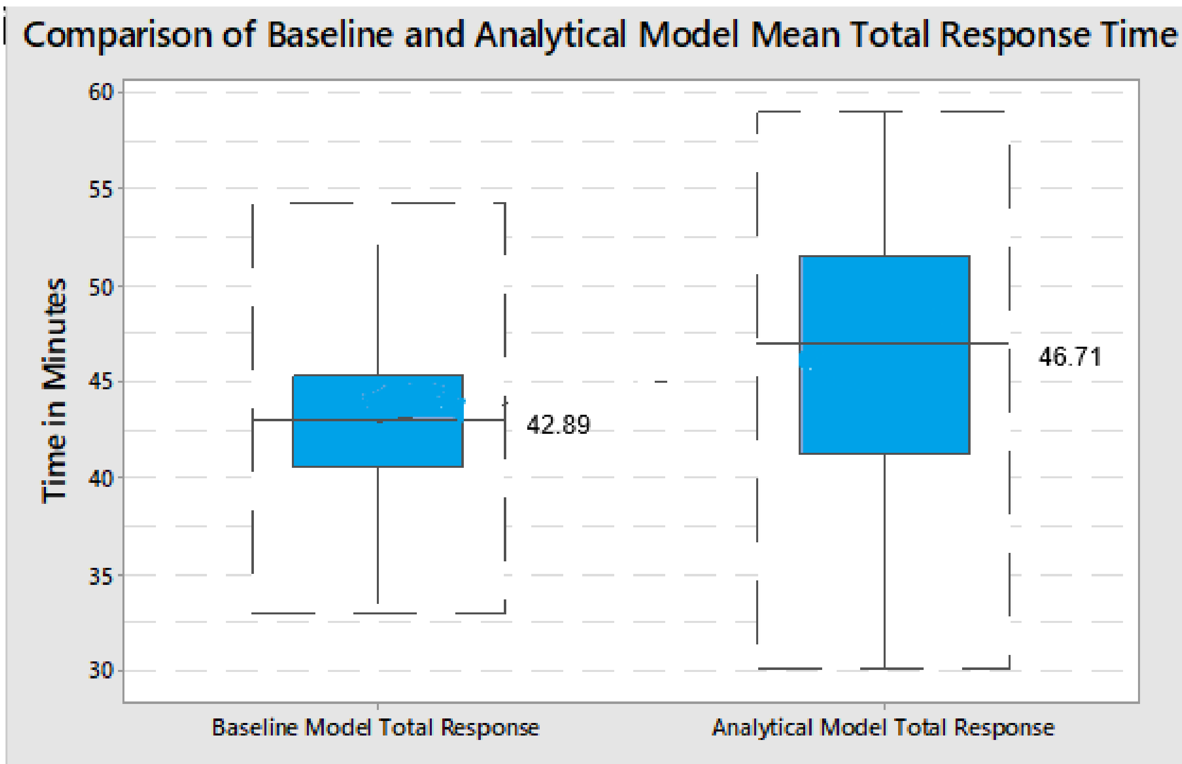
5.3 Crew Capability and Task Support Impacts

The results of section 5.1 and 5.2 provide evidence of the need for additional investigation in two areas:

1. Hypothesizing that the temporal response of sailors in the performance of their individual tasks is adversely affected by the assumed sailor capability and task support, and
2. Exploring the relationship between sailor capability and success rate in elimination of the small boat threat.

5.3.1 Evaluation of Temporal Response Impact

To confirm the hypothesis that the temporal response of sailors in the performance of their tasks is adversely impacted, and examine the magnitude of these impacts, the 1000 runs completed for both the baseline and analytical models were compared using Minitab statistical software. As noted earlier, in section 5.2, we identified that the individual temporal response times for each task in the analytical model were longer than the same tasks in the baseline model. This fact was the result of variability reduction and a weighting toward longer times resulting from the use of Equation (7). This same phenomenon is also demonstrated in the cumulative actions of the watch team and affects the total response time for the crew to eliminate the small boat threat as seen in Figure 24.



Variable	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Baseline Model Total Res	42.891	0.107	3.374	32.979	40.551	42.977	45.301	54.254
Analytical Model Total R	46.709	0.206	6.523	30.050	41.250	47.000	51.500	59.000

Figure 24: Boxplot Comparison of Baseline and Analytical Response Time

Similar to the computations completed to analyze individual means in Section 5.2, analysis of the total means allows the development of a conclusion regarding the hypothesis that using Equation (7) to include temporal variability impacts for sailor capability and task support results in a statistically significant time increase. Therefore, once again using a two-sample t-test to verify the null hypothesis that the means are equal results in a t-value of -16.44 and a P-Value of less than 0.001, allowing us to reject the null hypothesis and clearly identify that the model using Equation (7) results in longer mean times for both the individual tasks and the overall scenario total response completion time.

5.3.2 Evaluation of Sailor Capability Impact

As mentioned in Section 5.2, each individual watch team agent was assigned a likelihood of high capability and high task support, based on assumed shipboard profiles for crew rotation, crew qualification, and experience. Watch Officers were assigned a likelihood of high capability 50% of the time and other watch team members were assigned a 30% likelihood of high capability. These assignments were subjective based on the experience of the author, and the resultant impacts on the ability of the modeled crew to successfully eliminate the small boat warranted investigation. Thus, 1000 runs were completed for each value of sailor capability from 30-100% in 10% increments using Minitab statistical software. The cumulative performance of the watch team and their ability to successfully eliminate the small boat as a threat within the established 40 minute timeframe improved substantially as high capability likelihood increased, as seen in Figure 25.

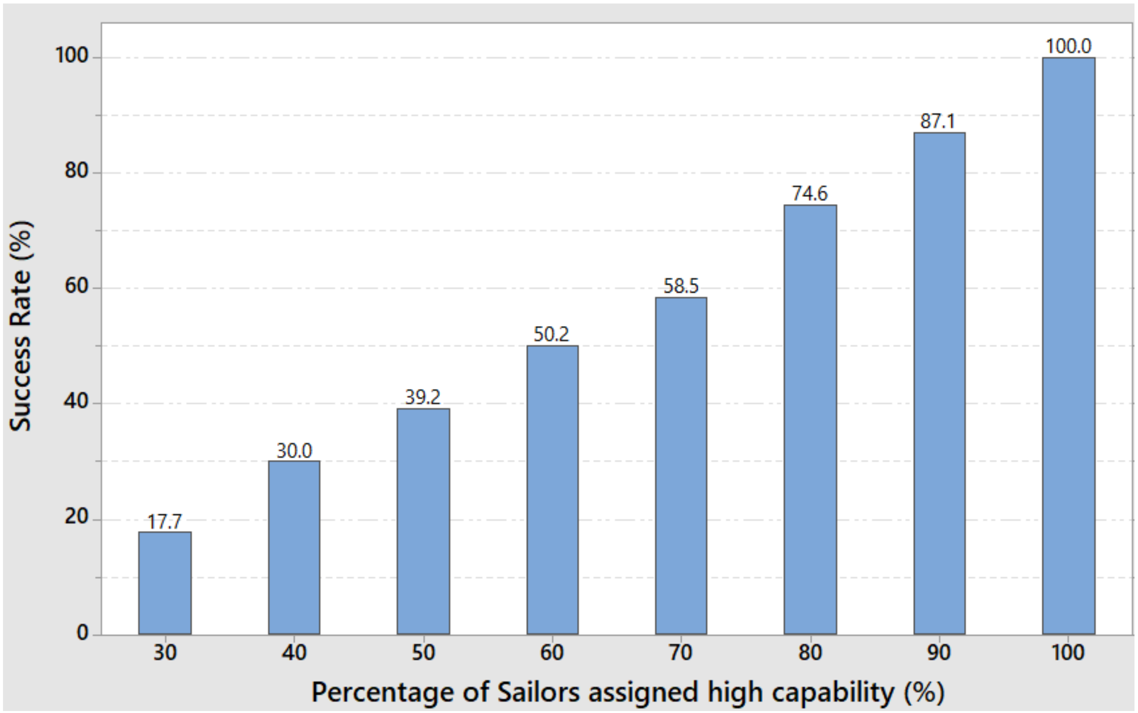


Figure 25: Comparison of Small Boat Elimination Success as a Function of Capability

These results indicate a linear relationship between sailor capability and the overall likelihood of success in eliminating the small boat as a threat within the required timeframe. They also establish a basis for an area of focus to mitigate risk to the ship. Recruiting, training, and assigning the most capable sailors to watch team positions evaluated in the modeled scenario significantly reduced the likelihood of negative consequences to the ship and its crew.

CHAPTER SIX: CONCLUSIONS

6.1 Summary

This research clearly identifies that analyses of temporal variation associated with any task execution requires consideration of many factors including the work pace and possible synergistic influence of actions on one another. It also demonstrates the feasibility of using a modularly built Agent Based Model to evaluate the impacts of task support and sailor capability on human temporal performance. And it clearly illustrates needed focus by the U.S. Navy to assure high levels of sailor capability for each of the small boat defense tasks in order to assure elimination of the small boat threat prior to the critical time.

The research also adds to the growing literature regarding the interplay between the physical and cognitive abilities of the individual in completing a given task and the impacts of resulting temporal variations. Variation occurs on both an intra- and inter-individual basis and is impacted by a wide variety of performance influencing factors (PIFs). Over the last three decades significant efforts have emerged to use, demonstrate, and apply a multitude of techniques to include Discrete Event Simulation, Bayesian Belief Networks, and Neural Networks, as well as a multitude of existing modeling software to provide relevant assessments of human task performance and temporal variability. Results have been applied to a wide range of socio-technical system applications with varying degrees of success. This study demonstrated ABMS as a method of assessing crew watch team response aboard U.S. Navy ships. As a test scenario, a group of sailors (agents) was assembled in an ABM to examine the task timelines and impact of temporal variability in crew performance. These simulations included human performance models for six crew members (agents) as well as a threat craft, and used models representing varying levels of crew capability and task support. In doing so, this study provided several conceptual developments in ABMS. Using AnyLogic, sophisticated human performance models were incorporated into the larger simulation of the entire shipboard system. Additionally, this work demonstrated a novel approach to using agent based models in an expanded environment for evaluation of task timelines and temporal variability impacts. These models were adapted and built to assure extensibility to support use across a broad range of U.S. Navy shipboard operations, using a series of agents within the simulation. The results of the experiment highlighted the ability of agent-based modeling and simulation to simultaneously provide detailed measures of individual sailor performance and of system-level emergent behavior. The individual measures of performance provide insight into the way the sailor will act within (and contribute to) the larger environment, and they reveal the demands of the larger environment on the individual watch team agent.

As research continues to mature in the area of human performance temporal variability, current momentum to move beyond singular and/or discrete applications of methods to assess human temporal variability must be maintained. Dependent upon the type of abilities (cognitive, psychomotor, physical) being exercised in the performance of a particular task, a variety of time distributions/equations can be integrated through multi-method modeling to provide far greater insight into sailor response as well as into the impacts and mitigation techniques for temporal variability in shipboard applications. This insight will lead to better understanding of the cumulative effects of time variability and will help to generalize and quantify taxonomies of influencing factors for use in both temporal performance and human reliability studies.

6.2 Recommendations and Future Research

Efforts to refine and further develop the model used in this analysis should continue. Sustained refinement should focus on improving data input methods for scenarios consisting of a large number of tasks. In addition, efforts to include a broader range of time distributions along with enhanced interactions between individual sailors should be considered. The assessment of additional factors that impact sailor temporal performance (i.e., task complexity) and a broader category of stochastic considerations (i.e., human response variability with different performance factors) would also improve the fidelity of the predictive outcome.

For the individual case of the Small Boat Defense Scenario, U.S. Navy shipboard leaders need to assess the relevance of the predictive outcomes from this ABMS methodology to their design, recruiting, and training processes. Improving this model with focused research, formally incorporating empirical data, and establishing operational decision making architectures will enable leadership to assess the need for task support improvements and qualification standards for assigned watch team members. The ability to integrate and examine decision-making activities with resultant error by including components of systems dynamics into the model would also add considerable value to this ABMS approach. These approaches to building upon and improving the use of ABMS in assessing shipboard operations would both reveal fundamental elements of agent behavior and provide greater insight into the impacts of the broader environment on the decision-making practices of the individual. Sustained and focused research on task influencing factors and their impacts needs to continue as well with the aim of optimizing a tailored PIF taxonomy for shipboard applications. The current practice of developing large and disparate taxonomies for each application presents unwieldy, unique, and niche frameworks for assessment of human temporal performance. A cost/benefit analysis of all the added detail must be completed to assure efficient and cost effective leadership decisions and sailor actions aboard every U.S. Navy vessel.

REFERENCES

Abbott, E. L., Liu, Z., Chua, D. K., & Chang, W. S. (2011). Estimating man-hours to construct a ship's block using work measurement. *Proceedings of the 2011 International Conference on Engineering, Project, and Production Management* (pp. 43-50). Singapore, Singapore: University of Singapore.

Ability[Def 2]. (n.d.). Retrieved May 2015, from Dictionary.com: http://dictionary.reference.com/browse/ability

Abrams, R. A., Meyer, D. E., & Kornblum, S. (1990). Eye-Hand Coordination: Oculomotor Control in Rapid Aimed Limb Movements. *Journal of Experimental Psychology: Human Perception and Performance, 16*(2), 248-267.

Aft, L. (2010). Don’t Abandon your work measurement cap. *Industrial Engineer*, 36–41.

Allender, L. (2000). Modeling Human Performance: Impacting system design, performance, and cost. In M. Chinni (Ed.), *Proceeding of the Military, Government, and Aerospace Simulation Symposium, 2000 Advanced Simulation Technologies Conference*, (pp. 139-144). Washington, DC.

Altmann, E. M., & Trafton, J. G. (2007). Timecourse of recovery from task interruption: Data and a model. *Psychonomic Bulletin and Review, 14*(6), 1079-1084.

Andreassi, J. L., & Huntley, M. S. (1967). *Effects of signal patterning upon vigilance performance and physiological responses.* Orlando, Florida: Naval Device Training Center.

Azarkhil, M., & Mosleh, A. (2014). Dynamic behavior of operating crew in complex systems. *The 60th Annual Reliability and Maintainability Symposium* (pp. 7-14). Colorado Springs, CO: Institute of Electrical and Electronics Engineers Inc. doi:10.1109/RAMS.2014.6798446

Badler, N. I., Albeck, J., Lee, S. J., Rabbitz, R. J., Broderick, T. T., & Mulkern, K. M. (2005). New behaviorial paradigms for virtual human models. *Proceedings of the 2005 DIgital Human Modeling for Design and Engineering Symposium* (pp. 43-50). Iowa City, IA: SAE International.

Balci, O. (. (1998). Verification, Validation, and Testing. In J. Banks, *Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice* (pp. 335 - 396). New York: Wiley.

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry: The power of response time distributional analyses. *Current Directions in Psychological Science*, 160-166.

Balota, D., & Spieler, D. (1999). Frequency, repetition, and lexicality effects in word recognition: Beyond measures of central tendency. *Journal of Experimental Psychology, 128*(32), 32-55.

Baqueiro, O., Wang, Y. J., McBurney, P., & Coenen, F. (2009). Integrating Data Mining and Agent Based Modleing and Simulation. *Advances in Data Mining: Applications and Theoretical Aspects - 9th Industrial Conference* (pp. 220-231). Leipzig, Germany: Springer.

Basahel, A. M., Young, M. S., & Ajovalasit, M. (2010). Effects of interaction between physical and mental workload on human performance. In M. Anderson (Ed.), *Contemporary Ergonomics and Human Factors 2010* (pp. 598-601). London, England: Taylor & Francis.

Beamish, S., Bhatti, S., Chubbs, C. S., MacKenzie, I. S., Wu, J., & Jing, Z. (2009). Estimation of psychomotor delay fro the Fitts' law coefficients. *Biological Cybernetics, 101*, 279-296.

Bedny, G. Z., & Karwowski, W. (2004). Activity theory as a basis for the study of work. *Ergonomics, 47*(2), 134–153. doi:10.1080/00140130310001617921

Bedny, G. Z., Karwowski, W., & Voskoboynikov, F. (2015). Applications of standardized motions in temporal analysis of work activity. *Human Factors and Ergonomics in Manufacturing and Service Industries, 25*(4), 469-483.

Bedny, G., & Karwowski, W. (2013). Analysis of strategies employed during upper extremity positioning actions. *Theoretical Issues in Ergonomics Science, 14*(2), 175–194. doi:10.1080/1463922X.2011.611265

Blackman, H. S., Gertman, D. I., & Boring, R. L. (2008). Human Error Quantification using Performance Shaping Factors in the SPAR-H Method. *Conference Proceedings for the 52nd Annual Meeting of the Human Factors and Ergonomics Society*, (pp. 1733-1737).

Bliss, J. P., & Chancey, E. T. (2014). An investigation of training strategies to improve alarm reactions. *Applied Ergonomics, 45*(5), 1278–1284. doi:doi.org/10.1016/j.apergo.2013.06.007

Bohannon, R. W., & Andrews, A. (2011). Normal walking speed: a descriptive meta-analysis. *Physiotherapy, 97*(3), 182–189. doi:10.1016/j.physio.2010.12.004

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America* (pp. 7280–7287). Washington, D.C.: National Academy of the Sciences Press.

Bootsma, R. J., Marteniuk, R. G., MacKenzie, C. L., & Zaal, F. T. (1994). The speed-accuracy trade-off in manual prehension: effects of movement amplitude, object size, and object width on kinematic characteristics. *Experimental Brain Research, 98*, 535-541.

Borah, J. (1995). *Investigation of Eye and Head Controlled Cursor Positioning Techniques.* Air Force Materiel Command, Armstrong Laboratory, Crew Systems Directorate. Wright-Patterson AFB, Ohio: Air Force Materiel Command. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA303807

Boring, R. (2010). *How many performance shaping factors are necessary for human reliability analysis?* Idaho Falls, ID: Idaho National Laboratory.

Bubb, H. (2005). Human Reliability: A key to improved quality in manufacturing. *Human Factors and Ergonomics in Manufacturing, 15*(4), 353-368.

Bustamante, E., Bliss, J., & Anderson, B. (2007). Effects of varying the threshold of alarm systems and workload on human performance. *Ergonomics, 50*(7), 1127-1147. doi:DOI: 10.1080/00140130701237345

Campbell, B. A., O'Brien, K. R., Byrne, M. D., & Bachman, B. J. (2008). Fitts' Law Predictions with an Alternative Pointing Device. *Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting* (pp. 1321-1325). New York, NY: Human Factors and Ergonomics Society.

Card, S. K., Moran, T. P., & Newell, A. (Eds.). (1986). *The Psychology of Human Computer Interaction.* Boca Raton, FL: CRC Press.

Carey, J. M., & Kacmar, C. J. (1997). The impact of communication mode and task complexity on small group performance and member satisfaction. *Computers in Human Behavior, 13*(1), 23–49. doi:10.1016/S0747-5632(96)00027-1

Cha, Y., & Myung, R. (2010). Extended Fitts' law in three-dimensional pointing tasks. *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting* (pp. 972-976). San Francisco, CA: Human Factors and Ergonomics Society.

Chan, A. H., Shum, V. W., Law, H. W., & Hui, I. K. (2003). Precise effects of control position, indicator type, and scale side on human performance. *International Journal of Advanced Manufacturing, 22,* 380-386. doi:10.1007/s00170-002-1491-z

Chang, Y., & Mosleh, A. (2007). Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents: Part 1-5. *Reliability Engineering & System Safety, 92*(8), 991-1101.

Chen, J. Y., & Joyner, C. T. (2006). Individual Differences in concurrent performance of Gunner's and robotic operator's tasks. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 1759-1763). San Francisco, CA: Human Factors and Ergonomics Society.

Cheng, R. (2011). Using Pearson Type IV and other Cinderella Distributions in Simulation. In S. Jain, R. Creasey, J. Himmelspach, K. White, & M. Fu (Ed.), *Proceedings of the 2011 Winter Simulation Conference* (pp. 457-468). Phoenix, AZ: Winter Simulation Conference.

Choodoung, S., & Smutkupt, U. (2012). Analysis Design for Assembly by MTM-2 (Methods Time Measurement-2) for Wood Joints in furniture. *Advanced Materials Research, 566*, 394-398.

Christmansson, M., Falck, A. C., Amprazis, J., Forsman, M., Rasmussen, L., & Kadefors, R. (2000). Modified method time measurements for ergonomic planning of production systems in the manufacturing industry. *International Journal of Production Research, 38*(17), 4051-4059. doi:10.1080/00207540050204911

Collins, S. H., & Kuo, A. D. (2013). Two Independent Contributions to Step Variability during Over-Ground Human Walking. *PLoS ONE, 8*(8), 1-11. doi:10.1371/journal.pone.0073597

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist, 12*, 671-684.

Crossman, E. R. (1959). A Theory of the Acquisition of Speed-Skill. *Ergonomics, 2*(2), 153-166.

Davidson, P. (2017). *Comprehensive Review of Recent Surface Force Incidents.* Norfolk, VA: U.S. Fleet Forces Command.

De Ambroggi, M., & Trucco, P. (2011). Modelling and assessment of dependent performance shaping factors through analytic network processes. *Reliability Engineering and System Safety, 96*, 849-860.

Department of Defense. (1989). Proceedings of a NATO Research Study Group 9 Workshop on Applications of Human Performance Models to System Design. In G. R. McMillan, D. Beevis, S. Eduardo, M. H. Strub, R. Sutton, & L. Van Breda (Ed.), *Defense Research Series. 2*, pp. 1-557. New York, New York: Plenum Press.

Department of Defense. (1997). *DoD Work Measurement / Labor Standards.* Defense Civilian Personnel Management Service, Defense Productivity Support Office. Arlington, VA: Defense Enterprise Integration Services.

Dietz, A. S., Weaver, S. J., Sierra, M. J., Bedwell, W. L., Salas, E., Fiore, S. M., . . . Driskell, J. E. (2010). Unpacking the Temporal and Interactive Effects of Stress on Individual and Team Performance. *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting*, (pp. 1017–1021).

Dixon, S. R., Wickens, C. D., & Chang, D. (2005). Mission Control of Multiple Unmanned Aerial Vehicles: A Workload Analysis. *Human Factors, 47*(3), 479-487.

Dodge, R. (1931). *Conditions and consequences of human variability.* New Haven, CT: Yale University Press.

Drewes, H. (2010). Only one Fitts' law formula please! *Proceedings of the Human Factors in Computing Sciences 28th Annual Meeting* (pp. 2813-2822). Paris, France: Association for Computing Machinery.

Druzdzel, M. J., & van der Gaag, L. C. (2000). Building probabilistic networks: "Where do the numbers come from?" Guest editors introduction. *IEEE Transactions on Knowledge and Data Engineering, 12*(4), 481-486.

Dumont, C., & Mazer, B. (2013). Assessment of Computer Task Performance for children: Standardization, validation and reliability. *Technology & Disability, 25*(1), 27-35. doi:10.3233/TAD-130364

Dunn, N., & Williamson, A. (2012). Mitigating the effects of Monotony:The role of task complexity. In *Rail human factors around the world: Impacts on and of people for successful rail operations* (pp. 774–782). CRC Press. doi:10.1002/lary.24911

El-Ladan, S. B., & Turan, O. (2012). Human Reliability Analysis-Taxonomy and praxes of human entropy boundary conditions for marine and offshore applications. *Reliability Engineering and System Safety, 98*, 43-54.

Eskenazi, T., Rotshtein, P., Grosjean, M., & Knoblich, G. (2012). The neural correlates of Fitts's law in action observation: An fMRI study. *Social Neuroscience, 7*(1), 30-41. doi:10.1080/17470919.2011.576871

Fereidunian, A., Zamani, M. A., Fatah, B., Lesani, H., Lucas, C., Kharzami, P., & Torabi, H. (2010). Investigation of Performance Shaping Factors in a Practical Human-Automation Interaction System. *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics* (pp. 1825-1831). Istanbul, Turkey: IEEE.

Fishwick, P. A. (1997). Computer Simulation: growth through extension. *Transactions of the society for Computer Simulation, 14*(1), 13-24.

Fitts, P. M. (1954). The information capacity of the human motor system in contolling the amplitude of movement. *Journal of Experimental Psychology, 47*, 381-391.

Fitts, P. M., & Peterson, J. R. (1964). Information capacity of discrete motor responses. *Journal of Experimental Psychology, 67*, 103-112.

Fleishman, E. A., & Quaintance, M. K. (1984). *Taxonomies of Human Performance: The Description of Human Task.* San Diego, CA: Academic Press.

Fletcher, A., & Dawson, D. (2001). A quantitative model of work-related fatigue: empirical evaluations. *Ergonomics, 44*(5), 475-488. doi:10.1080/00140130119824

Freivalds, A. (2009). *Niebel's Methods, Standard, and Work Design* (12th ed.). New York, NY: McGraw Hill.

Friedman Nir: Goldszmidt, M. (1996). Learning Bayesian networks with local structure. *Proceedings of the Twelfth International Conference on Uncertainty in Artificial Intelligence* (pp. 252-262). Berlin, Germany: Morgan Kaufman Publishers Inc.

Fuller, H. J., Reed, M. P., & Liu, Y. (2010). Integrating physical and cognitive human models to represent driving behavior. *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting* (pp. 1042-1046). San Francisco, CA: Human Factors and Ergonomics Society.

Gan, K. C., & Hoffman, E. R. (1988). Geometrical conditions for ballistic and visually controlled movements. *Ergonomics, 31*, 829-839.

Garrido, J. M. (2001). *Object-oriented discrete-event simulation with Java: a practical introduction.* New York, NY: Kluwer Academic/Plenum Publishers.

Giachetti, R. (2010). *Design of Enterprise Systems: Theory, Architecture, and Methods.* Boca Raton, FL, : CRC Press.

Gilbert, N. (2008). *Agent-based models.* Guilford, UK: Sage Publications.

Gillie, T., & Broadbent, D. (1989). What makes interruptins disruptive? A study of length, similarity, and complexity. *Psychological Research, 50*, 243-250.

Gore, B. F., Hooey, B. L., Foyle, D. C., & Scott-Nash, S. (2008). Meeting the challenge of cognitive human performance model interpretability through transparency: MIDAS v5.x. *Proceedings of the 2nd International Conference on Applied Human Factors and Ergonomics*. Las Vegas, NV.

Goure, D., & Ranney, J. M. (2000). *Averting the Defense Train Wreck in the New Millenium.* Washington, D.C.: Center for Strategic and International Studies.

Grabbe, J. W., & Allen, P. A. (2012). Cross-Task Compatibility and Age-Related Dual-Task Performance. *Experimental Aging Research, 38*(5), 469–487. doi:10.1080/0361073X.2012.726154

Gregoriades, A., & Sutcliffe, A. (2008). Workload prediction for improved design and reliability of complex systems. *Reliability Engineering and System Safety, 93*, 530-549.

Griffith, C. D., & Mahadevan, S. (2011). Inclusion of fatigue effect in human reliability analysis. *Reliability Engineering and System Safety, 96*, 1437-1447.

Grosjean, M., Shiffrar, M., & Knoblich, G. (2007). Fitts's law holds for action perception. *Psychological Science, 18*(2), 95-99.

Groth, K. M., & Mosleh, A. (2012). A data-informed PIF hierarchy for model-based human reliability analysis. *Reliability Engineering and System Safety, 108*, 154-174.

Guilford, J. P. (1956). The Structure of Intellect. *Psychological Bulletin, 53*, 267-293.

Guilford, J. P. (1959). Traits of Creativity. (H. H. Anderson, Ed.) *Creativity and its cultivation*, 142-161.

Hancock, P. A., & Caird, J. K. (1993). Experimental Evaluation of a model of mental workload. *Human Factors, 35*(3), 413-429.

Hancock, P. A., Sawyer, B. D., & Stafford, S. (2015). The effects of display size on performance. *Ergonomics, 58*(3), 337-354. doi:10.1080/00140139.2014.973914

Harman, E., Frykman, P., Pandorf, C., Tharion, W., Mello, R., Obusek, J., & Kirk, J. (1999). *Physiological Performance Comparisons of Soldiers Carrying Loads.* Fort Detrick, MD: US Army Medical Research and Materiel Command.

Hart, S. G. (2006). Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 50*, pp. 904-908. San Francisco, CA: SAGE Publications.

Hart, S. G., & Staveland, L. E. (1988). *Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research.* Retrieved from http://www.sciencedirect.com/science/article/pii/S0166411508623869

Heathcote, A., Popiel, S. J., & Mewhort, D. J. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin, 109*, 340-347.

Hick, W. E. (1952). On the rate of information gain. *Quarterly Journal of Experimental Psychology, 4*, 11-26.

Hocking, C., Silberstein, R. B., Lau, W. M., Stough, C., & Roberts, W. (2001). Evaluation of cognitive performance in the heat by functional brain imaging and psychometric testing. *Comparative Biochemistry and Physiology, 128*, 719-734.

Hoffmann, E. R. (2013). Which version/variation of Fitts" law? A critique of Information Theory Models. *Journal of Motor Behavior, 45*, 205-215.

Hollnagel, E. (1993). *Human Reliability Analysis: Context and Control.* London: Academic Press.

Huang, J. X., Bian, Y. Q., & Li, S. B. (2013). Research of equipments maintenance support analysis based on the expanded ESD. *Applied Mechanics and Materials, 419*, 820-825.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology, 45*(3), 188-196.

Itami, A. S., Antonio, F. J., & Mendes, R. S. (2015). Very prolonged practice in block of trials: Scaling of fitness. *Physica A: Statistical Mechanics and its applications, 424*, 82-89.

Jennings, N. (2000). On Agent Based Software Engineering. *Artificial Intelligence, 117*(2), 277-296.

John, B. E., Patton, E. W., Gray, W. D., & Morrison, D. F. (2012). Tools for Predicting the Duration and Variability of Skilled: Performance without Skilled Performers. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 56*, pp. 985–989. Boston, Massachusetts: Sage Publications. doi:10.1177/1071181312561206

Kamienkowski, J., & Sigman, M. (2008). Delays without Mistakes: Response Time and Error Distributions in Dual-Task. *PLoS ONE, 3*(9), 1-7. doi:doi:10.1371/ journal.pone.0003196

Karwowski, W. (2003). Considering the importance of individual differences in human factors research: No longer simply confounding noise. Panel Summary. *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting* (pp. 1082-1086). Santa Monica, CA: Human Factors and Ergonomics Society.

Karwowski, W. (2012). A review of human factors challenges of complex adaptive systems: discovering and understanding chaos in human performance. *Human Factors, 54*, 983-995. doi:10.1177/0018720812467459

Karwowski, W. (2013). *Modeling the ideal human observer concept.* Orlando, FL: University of Central Florida.

Karwowski, W., Kern, D., Muhs, K. S., Dumond, D., Skahen, D., & Freeman, J. (2014). Modeling Human Execution TImes in STAMPS. *Unpublished Work*. Orlando, FL.

Kern, D. J., & Culley, K. E. (2015). How do Submarine Force trainers use computer-based navigation trainer systems? Applying cluster analysis to examine situated objectives in the employment of adaptive CBT. *Computers in Human Behavior, 49*, 313-323.

Kim, J. W., & Jung, W. (2003). A taxonomy of performance influencing factors for human reliability analysis of emergency tasks. *Journal of Loss Prevention in the Process Industries*, *16*, 479-495.

Kobus, D. A., Brown, C. M., Wu, L., Robusto, K., & Bartlett, J. (2011). *Cognitive performance and physiological changes under heavy load carriage.* San Diego, CA: Pacific Science and Engineering Group.

Koh, R. Y., Park, T., & Wickens, C. D. (2014). An investigation of differing levels of experience and indices of task management in relation to scrub nurses' performance in the operating theatre: Analysis of video-taped caesarean section surgeries. *International Journal of Nursing Studies*, *51*(9), 1230-1240. doi:10.1016/j.ijnurstu.2014.01.005

Kuhlang, P., Edtmayr, T., & Sihn, W. (2011). Methodical approach to increase productivity and reduce lead time in assembly and production-logistic processes. *CIRP Journal of Manufacturing Science and Technology, 4,* 24-32.

Kuo, C. F., & Wang, M. J. (2009). Motion generation from MTM semantics. *Computers in Industry, 60*, 339-348.

Laring, J., Forsman, M., Kadefors, R., & Örtengren, R. (2002). MTM-based ergonomic workload analysis. *International Journal of Industrial Ergonomics, 30*(3), 135–148. doi:10.1016/S0169-8141(02)00091-4

Lee, S. W., Kim, A. R., Ha, J. S., & Seong, P. H. (2011). Development of a qualitative evaluation framework. *Annals of Nuclear Energy, 38*, 1751-1759.

Lee, S., Heo, G., & Chang, S. H. (2006). Prediction of the human response time with the similarity and quantity of information. *Reliability Engineering and System Safety, 91*, 728-734.

Lee, T. J., Tseng, S. S., & Kung, C. C. (2015). Factors Influencing the Performance of Human Computation : An Empirical Study in Web Application Testing. *Journal of Information Science and Engineering, 31*, 641–658.

Lehto, M. R., & Landry, S. (2012). *Introduction to Human Factors and Ergonomics for Engineers.* Boca Raton, FL: CRC Press.

Lemmer, J., & Gossink, D. (2004). Recursive noisy OR-A rule for estimating complex probabilistic interactions. *IEEE Transactions of Systems, Man, and Cybernetics Part B, 34*(6), 2252-2261.

Li, Y. (2013). *Modeling and simulation of operator knowledge-based behavior (Doctoral Dissertation).* Ann Arbor, MI: ProQuest LLC.

Limpert, E., Stahel, W. A., & Abbt, M. (2001). Log-normal Distributions across the Sciences: Keys and Clues. *BioScience, 51*(5), 341-352.

Lin, C. J., Jou, Y. T., Yenn, T. C., Hsieh, T. L., & Yang, C. W. (2009). A study of control room staffing and workload from the human information processing perspective. *IEEE International Conference on Networking, Sensing and Control* (pp. 303-307). Okayama, Japan: IEEE.

Lin, C.-J., & Wu, C. (2013). Reactions, accuracy and response complexity of numerical typing on touch screens. . *Ergonomics, 56*(5), 818-831. doi:10.1080/00140139.2013.767384

Lin, M. L., Radwin, R. G., & Vanderheiden, G. C. (1992). Gain effects on performance using a head-controlled computer input device. *Ergonomics, 35*(2), 159-175.

Liu, F. C. (2011). Validation And Agent-Based Modeling: A Practice Of Contrasting Simulation Results With Empirical Data. *New Mathematics and Natural Computation, 7*(3), 515-542.

Luce, R. D. (1986). *Response Times: Their role in inferring elementary mental organization.* New York, NY: Oxford University Press.

Macal, C., & North, M. (2014). Introductory Tutorial: Agent-Based Modeling and Simulation. *Proceedings of the 2014 Winter Simulation Conference* (pp. 6-20). Savannah, GA: IEEE Press.

Mackieh, A., & Cilingir, C. (1998). Effects of performance shaping factors on human error. *International Journal of Industrial Ergonomics, 22,* 285–292. doi:10.1016/S0169-8141(97)00079-6

Madden, D. J., Gottlob, L. R., Denny, L. L., Turkington, T. G., Provenzale, J. M., Hawk, T. C., & Coleman, R. E. (1999). Aging and recognition memeory: Changes in regional cerebral blood flow associated with components of reaction time distributions. *Journal of Cognitive Neuroscience, 11*(5), 511-520.

Mandal, S., Han, X., Pattipati, K., & Kleinman, D. (2010). Agent-Based Ditributed Framework for Collaborative Planning. *Proceedings of the 2010 IEEE Aerospace Conference* (pp. 1-11). Big Sky, MT: IEEE .

Marras, W. S., & Hancock, P. A. (2014). Putting mind and body back together: A human systems approach to the integration of the physical and cognitive dimennsions of task design and operations. *Applied Ergonomics, 45*(1), 55-60.

Martin, F. M. (2009). A theory of reaction time distributions. *Unpublished Work. University of Provence*. Marseilles, France.

Marusich, L. R., Buchler, N., & Bakdash, J. Z. (2014). *Human limits to cognitive information fusion in a military decision-making task.* Alexandria, VA: Defense Technical Information Center.

Mason, S., Baines, T., Kay, J. M., & Ladbrook, J. (2005). Improving the design process for factories: Modeling human performance variation. *Journal of Manufacturing Systems, 24*(1), 47–54. doi:10.1016/S0278-6125(05)80006-8

Matthews, G., Warm, J. S., Dember, W. M., Mizoguchi, H., & Smith, A. P. (2001). The common cold impairs visual attention, psychomotor performance, and task engagement. *Proceeding of the*

Human Factors and Ergonomics Society 45th Annual Meeting (pp. 1377-1381). Minneapolis, MN: Human Factors and Ergonomics Society.

Matzke, D., & Wagenmakers, E. J. (2009). Psychological interpretation of the ex-Gaussian and Shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review, 16*(5), 798-817.

Maynard, H. B., Stegemerten, G. J., & Schwab, J. L. (1948). *Methods-time Measurement.* New York, NY: McGraw-Hill.

Meirina, C., Levchuk, G., & Pattipati, K. (2003). A multi-agent decision framework for DDD-III environments. *Proceedings of the 8th International Command and Control Research and Technology Symposium.* Washington, D.C.: Naval Postgraduate School.

Mellouli, S., Moulin, N. B., & Mineau, G. (2004). Laying down the foundations of an agent modelling methodology for fault-tolerant multi-agent systems. *Engineering Societies in the Agents World IV*, 275-293. Springer.

Meyers, F. E., & Stewart, J. R. (2002). *Motion and time study for lean manufacturing.* Upper Saddle River, NJ: Prentice Hall.

Micalizzi, J., & Wickens, C. D. (1981). *The Application of Additive Factors Methodology to Workload Assessment in a Dynamic System Monitoring Task.* Urban-Champaign, Illinois: Defense Technical Information Center.

Miles, J. D., & Proctor, R. W. (2012). Correlations between spatial compatibility effects: are arrows more like locations or words? *Psychological Research, 76*(6), 777-791. doi:10.1007/s00426-011-0378-8

Mindock, J. A., & Klaus, D. M. (2014). Contributing factor map: A taxonomy of influences on human performance and health in space. *IEEE Transactions on Human Machine Systems, 44*, 591-602.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analysis: The PRISMA Statement. *PLoS Med, 6*(7). doi:10.137/journal.pmed1000097

Mosleh, A., & Chang, Y. H. (2004). Model based human reliability analysis: Prospects and Requirements. *Reliability Engineering and System Safety, 83*, 241-253.

Mracek, D. L., Arsenault, M. L., Day, E. A., Hardy, J. H., & Terry, R. A. (2014). A Multilevel Approach to Relating Subjective Workload to Performance After Shifts in Task Demand. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 56*(8), 1401 –1413. doi:10.1177/0018720814533964

Muhs, K. S., Karwowski, W., & Kern, D. (2018). Temporal variability in human performance: A systematic literature review. *International Journal of Industrial Ergonomics, 64*, 31-50.

Murthy, U. S., & Kerr, D. S. (2003). Decision making performance of interacting groups: An experimental investigation of the effects of task type and communication mode. *Information and Management, 40*(5), 351–360. doi:10.1016/S0378-7206(02)00017-4

Nakayama, S., Nakayama, K., & Nakayama, H. (2001). A study on setting standard time using work achievement quotient. *International Journal of Production Research, 40*(15), 3945-3953. doi:10.1080/00207540210159581

North, M. J., & Macal, C. M. (2007). *Managing Business Complexity: Discovering Strategic Solutions with Agent Based Modeling and Simulation.* USA: Oxford University Press.

O'Boyle, E., & Aguinis, H. (2012). The Best & the Rest: Revisiting the Norm of Normality of Individual Performance. *Personnel Psychology, 65*, 79-119.

Olivier, J., & Norberg, M. M. (2010). Positively skewed data: Revisiting the Box-Cox power transformation. *International Journal of Psychological Research, 3*(1), 69-78.

O'Neal, E. K., & Bishop, P. (2010). Effects of work in a hot environment on repeated performances of multiple types of simple mental tasks. *International Journal of Industrial Ergonomics, 40*, 77-81.

Onisko, A., Druzdzel, M. J., & Wasyluk, H. (2000). Learning Bayesian network parameters from small data sets: Application of noisy-OR gates. *12th European Conference on artificial intelligence (ECAI-2000).* Berlin, Germany.

Onisko, A., Druzdzel, M., & Wasyluk, H. (2001). Learning Bayesian network parameters from small data sets; application of noisy-OR gates. *International Journal of Approximate Reasoning, 27*(2), 165-182.

Palmer, E. M., Horowitz, T. S., Torralba, A., & Wolfe, J. M. (2011). What are the shapes of response time distributions in visual search. *Journal of Experimental Psychology: Human Perception and Performance, 37*(1), 58-71.

Pearl, J. (1988). *Probabilisitic Reasoning in Intelligent Systems.* San Francisco, California: Morgan Kaufmann.

Perry, C. M., Sheik-Nainar, M. A., Segall, N., Ma, R., & Kaber, D. B. (2008). Effects of physical workload on cognitive task performance and situation awareness. *Theoretical Issues in Ergonomic Science, 9*(2), 95-113.

Podofillini, L., Park, J., & Dang, V. N. (2013). Measuring the influence of task complexity on human error probability: An empirical evaluation. *Nuclear Engineering and Technology, 45*(2), 151–164. doi:10.5516/NET.04.2013.702

Proctor, R. W., & Vu, K.-P. L. (2006). The Cognitive Revolution at Age 50: Has the Promise of the Human Information-Processing Approach Been Fulfilled? *International Journal of Human-Computer Interaction, 21*(3), 253-284.

Pyy, P. (2000). An approach for assessing human decision reliability. *Reliability Engineering & System Safety, 68*(1), 17-28. doi:10.1016/S0951-8320(99)00078-2

Rantanen, E. M., & Levinthal, B. R. (2005). Time-based modeling of human performance. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting* (pp. 1200-1204). Orlando, FL: Human Factors and Ergonomics Society.

Rasmussen, J. (1983). Skills, rules, and knowledge; Signals, signs, and symbols, and other distinctions in human performance models. *IEEE transactions on systems, man, and, SMC 13*, 257-266.

Rasmussen, J. (1997). Risk management in a dynamic society: A modeling problem. *Safety Science, 27,* 183-213.

Rasmussen, J. (1997). Risk Management in a dynamic society: A modelling problem. *Safety Science, 27*, 183-213.

Ratcliff, R., & Murdock, B. B. (1976). Retrieval processes in recognition memory. *Psychological Review, 83*, 190-214.

Razmi, J., & Shakhs-Niyaee, M. (2008). Developing a specific predetermined time study approach: An empirical study in a car industry. *Production Planning & Control: The Management of Operations, 19*(5), 454-460.

Reer, B. (1994). Probabilistic method for analyzing the reliability effect of time and organizational factors. *European Journal of Operational Research, 75*, 521-539.

Roth, E. M., Patterson, E. S., & Mumaw, R. J. (2001). Cognitive engineering: Issues in user-centered system design. In J. Marciniak (Ed.), *Encyclopedia of Software Engineering, 2nd Edition* (pp. 2042-2076). New York, New York: Wiley Interscience, John Wiley & Sons.

Roth, J. T. (1992). *Reliability and validity assessment of a taxonomy for predicting relative stressor effect on human task performance.* Boulder, CO: Micro Analysis and Design Inc.

Rouder, J. N., Lu, J., Speckman, P., Sun, D., & Jiang, Y. (2005). A hierarchal model for estimating response time distributions. *Psychonomic Bulletin & Review, 12*(2), 195-223.

Sabaric, I., Brnada, S., & Kovacevic, S. (2013). Application of the MODAPTS method with innovative solutions in the warping process. *Fibres and Textiles in Eastern Europe, 21*(4), 55-59.

Salvendy, G. (Ed.). (1997). *Handbook of Human Factors and Ergonomics.* New York, NY: John Wiley & Sons.

Salvendy, G. (Ed.). (2001). *Handbook of Industrial Engineering: Technology and Operations Management.* New York, NY: John Wiley & Sons.

Sargent, R. G. (2007). Simulation Model Verification and Validation. *Proceedings of the 2007 Winter Simulation Conference* (pp. 124-137). Washington, D.C.: IEEE.

Sasou, K., & Reason, J. (1999). Team Errors: Definition and Taxonomy. *Reliability Engineering and System Safety, 65*, 1-9.

Seow, S. C. (2005). Information theoretic models of HCI: A comparison of the Hick-Hyman law and Fitts' las. *Human Computer Interaction, 20*(3), 315-352.

Sheridan, T. B. (2013). Human Response is Lognormal; Plan on waiting if you want reliability. *Ergonomics in Design: The Quarterly of Human Factors Applications, 21*(4), 4-6.

Shi, J., & Shi, S. (2013). Various Degrees of Mental Workload and its Effect on Driver Reaction Time. *Proceedings of the Fourth International Conference on Transportation Engineering* (pp. 277–286). Chegdu, China: American Society of Civil Engineers.

Siebers, P. O., & Aickelin, U. (2008). Introduction to Multi-Agent Simulation. In F. Adam, & P. (. Humphreys, *Encyclopedia of Decision Making and Decision Support Technologies* (pp. 554-564). Hershey, Pennsylvania: Idea Group Publishing.

Skewness. (2014, May). Retrieved 2014, from Wikipedia: http//en.wikipedia.org/wiki/Skewness

Smith, T. J., Henning, R. A., Wade, M. G., & Fisher, T. (2014). *Variability in Human Performance.* Boca Raton, FL: CRC Press.

Song, H. T., Clawson, J., & Radu, I. (2012). Updating Fitts' law to account for small targets. *International Journal of Human-Computer Interaction, 28*, 433-444.

Speier, C., Vessey, I., & Valacich, J. S. (2003). The effects of interruptions, task complexity, and information presentation on computer-supported decision-making performance. *Decision Science, 34*, 771-797.

Stanton, N. A., & Baber, C. (2008). Modelling of human alarm handling response times: a case study of the Ladbroke Grove rail accident in the UK. *Ergonomics, 51*(4), 423-440. doi: 10.1080/00140130701695419

Stary, C., & Peschl, M. F. (1998). Representation still matters: Cognitive engineering and user interface design. *Behaviour and Information Technology, 17*(6), 338-360. doi:10.1080/014492998119292

Stern, J. A., & Brown, T. B. (2005). *Real-Time detector of human fatigue.* Saint Louis, MO: Bio-behavior Analysis Systems, LLC.

Sugarman, R. C. (2011). Something Old, Something New, Something Borrowed, but Nothing Blew! . *Cognitive Ergonomics*, 24-28.

Swain, A., & Guttman, H. (1983). *Handbook of Human Reliability analysis with emphasis on nuclear power plants* . Washington, DC: US Nuclear Regulatory Commission.

Teichner, W. H., Williams, E., Ekel, G., & Corso, G. M. (1979). *Information Processing Theory of Human Performance and Related Research.* Bolling Air Force Base, Washington DC: Air Force Research Laboratory.

Thurstone, L. L. (1938). *Primary Mental Abilities.* Chicago, IL: University of Chicago Press.

Tiwari, A. (2008). Small Boat and Swarm Defense: A Gap Study (Master's Thesis). Monterey, California, USA: Naval Postgraduate School.

Ulrich, R., & Miller, J. (1993). Information processing models generating lognormally distributed reaction times. *Journal of Mathematical Psychology, 37*, 513-525.

van der Linden, W. J. (2006). A lognormal model for response times on test items. *Journal of Educational and Behavioral Statistics, 31*(2), 181-204.

Van Zandt, T. (2000). How to fit a response time distribution. *Psychonomic Bulletin & Review, 7*(3), 424-465. doi:10.3758/BF03214357

Van Zandt, T. (2002). Analysis of Response Time Distributions. (J. W. Pashler, Ed.) *Stevens' handbook of experimental psychology, volume 4: Methodology in experimental psychology*, 461-516.

VanRullen, R., & Thorpe, S. J. (2001). The time course of visual processing: from early perception to decision-making. *Journal of Cognitive Neuroscience, 13*(4), 454-461. Retrieved from www.ncbi.nlm.nih.gov/pubmed/11388919

Waganaar, W. A., & Sagaria, S. D. (1975). Misperception of exponential growth. *Perception and Psychophysics, 18,* 416-422.

Waters, A. J. (1998). The Effects of Smoking on Performance on the Garner Speeded Classification Task. *Human Psychopharmacology Clinical and Experimental, 13*, 477-491.

Weaver, S. M., Foxe, J. J., Shpaner, M., & Wylie, G. R. (2006, June). You can’t always get what you want: The influence of unexpected task constraint on voluntary task switching. *Quarterly Journal of Experimental Psychology*, 1–13. doi:10.1080/17470218.2014.917115

Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record, 58,* 475-482.

Wickens, C. D. (1992). *Engineering Psychology and Human Performance.* New York, NY: Harper Collins.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Williams, C. (2013). Accepting Austerity: The Right Way to Cut Defense. *Foreign Affairs, 92*(6), 54-64.

Williams, J. (1988). A data-based method for assessing and reducing human error to improve operational performance. *4th IEEE conference on Human Factors in Nuclear Power Plants* (pp. 436-450). Monterey, CA: IEEE.

Wong, D., Walters, B., & Fairey, L. (2010). Validating human performance models of the future Orion crew exploration vehicle. *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting* (pp. 1042-1046). San Francisco, CA: Human Factors and Ergonomics Society.

Wooldridge, M. (2009). *An Introduction to MultiAgent Systems.* Chichester, UK: John Wiley & Sons Ltd.

Xiang, X., Kennedy, R., Madey, G., & Cabaniss, S. (2005). Verification and validation of agent-based scientific simulation models. *Proceedings of the 2005 Agent-Directed Simulation Symposium* (pp. 47-55). San Diego, CA: The Society for Modeling and Simulation International.

Yadav, T. K. (2013). Measurement Time Method for engine assembly with help of Maynard Operating Sequencing Technique (MOST). *International Journal of Innovations in Engineering and Technology, 2*, 131-136.

Yogev-Seligmann, G., Hausdorff, J. M., & Giladi, N. (2008). The role of executive function and attention in gait. *Movement Disorders, 23*(3), 329–342. doi:10.1002/mds.21720

Young, M. S., Brookhuis, K. A., Wickens, C. D., & Hancock, P. A. (2015). State of science: mental workload in ergonomics. *Ergonomics, 58*(1), 1-17.

Zagorecki, A., & Druzdzel, M. J. (2004). An empirical study of probability elicitation under Noisy-OR assumption. In V. Barr, & Z. Markov (Ed.), *Proceedings of the Seventeenth International Florida*

Artificial Intelligence Research Society Conference (FLAIRS -2004) (pp. 880-885). Miami Beach, FL: AAA Press.

Zhang, L. (2003). Cognitive task network modeling and human task simulation visualization in a 3-d environment (Preliminary feasibility study). *Recent Developments in Automotive Safety Technology*, 635-639.

Zheng, H., Son, Y.-J., Chiu, Y.-C., Head, L., Feng, Y., Xi, H., . . . Hickman, M. (2013). *A Primer for Agent-Based Simulation and Modeling in Transportation Applications.* McLean, VA: Office of Safety Research and Development.