NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

AN EVALUATION OF THE EFFECTIVENESS OF U.S. NAVAL AVIATION CREW RESOURCE MANAGEMENT TRAINING PROGRAMS: A REASSESSMENT FOR THE TWENTY-FIRST CENTURY OPERATING ENVIRONMENT

by

Douglas W. Jones

June 2009

Thesis Advisor: Paul E. O'Connor
Second Reader: Michael E. McCauley

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** June 2009	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE An Evaluation of the Effectiveness of U.S. Naval Aviation Crew Resource Management Training Programs: A Reassessment for the Twenty-First Century Operating Environment			**5. FUNDING NUMBERS**
6. AUTHOR(S) Douglas W. Jones			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited			**12b. DISTRIBUTION CODE**

13. ABSTRACT (maximum 200 words)

This thesis describes a multi-faceted evaluation of the U.S. Naval Aviation Crew Resource Management (CRM) program. CRM training is used to instruct naval aviators in safety critical, non-technical behaviors. Reactions were evaluated by using a single item from command safety climate questionnaires (n=51, 570 observations over nine years). Attitudes were assessed using a 37-item survey (364 responses). Knowledge was evaluated using a 10-item multiple-choice test (123 responses). Finally, the causes of naval aviation mishaps from fiscal years 1997-2007 (238 mishaps) were examined to identify how many were attributed to failings related to CRM concepts. It was found that aviators perceived CRM training to be useful, had positive attitudes towards concepts addressed in the training, and the level of knowledge was constant across rank and aircraft type. Nevertheless, human error still accounts for more than 80% of all mishaps in naval aviation, and over 65% of those are attributed to at least one failure in CRM. As human error continues to plague naval aviation, routine evaluations of CRM's effectiveness are critical to ensure it is achieving its goal to "improve mission effectiveness by minimizing crew preventable errors, maximizing crew coordination, and optimizing risk management" (CNO, 2001).

14. SUBJECT TERMS Crew Resource Management, CRM, Naval Aviation, Training Evaluation, Cockpit Resource Management, Safety, Aviation Mishaps, Training Effectiveness Evaluation			**15. NUMBER OF PAGES** 127
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; public distribution is unlimited

AN EVALUATION OF THE EFFECTIVENESS OF U.S. NAVAL AVIATION CREW RESOURCE MANAGEMENT TRAINING PROGRAMS: A REASSESSMENT FOR THE TWENTY-FIRST CENTURY OPERATING ENVIRONMENT

Douglas W. Jones
Lieutenant Commander, United States Navy
B.S., Westmont College, 1998

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN HUMAN SYSTEMS INTEGRATION

from the

NAVAL POSTGRADUATE SCHOOL
June 2009

Author: Douglas W. Jones

Approved by: Paul E. O'Connor, PhD
Thesis Advisor

Michael E. McCauley, PhD
Second Reader

Robert F. Dell, PhD
Chairman, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This thesis describes a multi-faceted evaluation of the U.S. Naval Aviation Crew Resource Management (CRM) program. CRM training is used to instruct naval aviators in safety critical, non-technical behaviors. Reactions were evaluated by using a single item from command safety climate questionnaires (n=51, 570 observations over nine years). Attitudes were assessed using a 37-item survey (364 responses). Knowledge was evaluated using a 10-item multiple-choice test (123 responses). Finally, the causes of naval aviation mishaps from fiscal years 1997-2007 (238 mishaps) were examined to identify how many were attributed to failings related to CRM concepts. It was found that aviators perceived CRM training to be useful, had positive attitudes towards concepts addressed in the training, and the level of knowledge was constant across rank and aircraft type. Nevertheless, human error still accounts for more than 80% of all mishaps in naval aviation, and over 65% of those are attributed to at least one failure in CRM. As human error continues to plague naval aviation, routine evaluations of CRM's effectiveness are critical to ensure it is achieving its goal to "improve mission effectiveness by minimizing crew preventable errors, maximizing crew coordination, and optimizing risk management" (CNO, 2001).

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

Human error is attributed as a causal factor in more than 80% of all mishaps in Naval aviation (Naval Safety Center, 2006). The Naval Aviation Crew Resource Management (CRM) Training Program was initiated by the Naval Safety Center in 1989 to reduce the number of aviation mishaps by minimizing the number of human performance-related errors (Alkov, 1989). CRM seeks to employ all available resources including human, informational and equipment to maximize performance while ensuring mission safety (Federal Aviation Administration, 2004). In 1991, the Department of the Navy officially adopted the practice of routinely administering CRM training to all aeronautically designated personnel (Chief of Naval Operations, 2001). Despite the numerous advancements over the last two decades in aircraft system design, minimal changes have been made to the Navy's CRM training program. Since the majority of CRM evaluations were performed over a decade ago, a need existed for additional research into the effectiveness of the current Naval CRM program. This thesis provides a thorough analysis of the Navy's CRM program, seeking to determine if the training continues to accomplish its designed objective of improving safety and crew performance.

The goal of this study was to evaluate the effectiveness of the Naval CRM training program using a multi-faceted assessment across different aviation platforms (Navy and Marine Corps). In addition, this evaluation considers the Human Systems Integration (HSI) approach to CRM by addressing human factors, safety, as well as manpower, personnel, training and education (MPT&E) strategies related to competency management and development within each aviation community. Kirkpatrick's (1976) hierarchy of training evaluation technique was applied to examine three levels of training assessment including: participant reactions, amount of learning received, and impact on the Navy as an organization (in terms of mishaps).

Reactions. Naval aviators CRM reaction responses indicated on command safety climate surveys were extracted from the Navy's Command Safety Assessment Survey (CSAS) from fiscal years 2000-2008. Over 50,000 responses were analyzed in terms of the perceived utility of Naval CRM training. The results suggest that aviators continue to be favorable about CRM training and perceive it as “useful.”

Learning. Attitudes and knowledge assessments were distributed to naval aviators to examine effects of learning with respect to non-technical concepts taught in CRM training. A Naval Aviator Human Factors (NAHF) questionnaire was developed and administered to 364 aviation officers investigating attitudes towards CRM concepts. Attitudes of naval aviators were generally positive, indicating acceptance of the CRM program. To further assess learning, a Knowledge Assessment for Naval Aviation CRM (KANCRM) was created and disseminated to 100 aviators to analyze differences between types of aviators in understanding CRM techniques. Despite lower than expected scores, no significant differences were observed between aviators varying in rank and airframe experience.

Results. The organizational impact of CRM was evaluated by examining aviation mishaps from 1997 to 2007. It was found that mishap rates have remained fairly consistent over the past several decades. The percentage of reported mishaps as a result of human error or CRM failure has remained over 65%. These findings suggest human error still plagues naval aviation despite the indicated cultural acceptance of CRM from the reactions and attitudes of naval aviators.

In conclusion, the results may imply that further reduction of human error in aviation has asymptotically reached a limit. Human error may very well be inevitable (Helmreich, Merritt, & Wilhelm, 1999). Due to the complexity and number of factors involved in aviation mishaps, estimating the success of CRM may be too challenging. The possibility exists that CRM is maintaining already low mishap rates and preventing them from rising to their previously elevated levels. Nonetheless, the findings of this evaluation have implications for human factors issues addressed in CRM training and highlight potential areas for future improvement of the Naval Aviation CRM program. Continuous assessment of Naval CRM training is essential in measuring the vitality of the program and whether it is meeting its goal in aviation of “improving mission effectiveness by minimizing crew preventable errors, maximizing crew coordination, and optimizing risk management” (Chief of Naval Operations, 2001).

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to several individuals involved in this project. First, I would like to thank Lieutenant Commander Paul O'Connor, thesis advisor, and Dr. Michael McCauley, my second reader. These dedicated professionals were instrumental in guiding and supporting the intensive research involved in this study. Without their assistance, this thesis effort would not have been possible. I would also like to extend my thanks to Dr. M. Quinn Kennedy and Dr. Ron Fricker for their assistance in providing valuable insight and experience in the analysis of my research. I extend my appreciation to various members of the Naval Safety Center in Norfolk, VA and Naval Aviation Schools Command in Pensacola, FL. Their willingness to coordinate efforts in data collection was essential in the timely completion of the thesis. Thanks to all the participants and aviators who took the time to provide feedback and valuable information that may influence any future improvements to CRM training. Finally, I would like to thank my wife, Dana, for being unconditionally supportive and enduring the long hours I spent away from my beloved family. I am grateful and thankful for all your help.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. HUMAN ERROR IN AVIATION

Concern for safety has remained a high visibility issue in both commercial and military aviation. A shift has already occurred; pinpointing the preponderance of aviation mishaps as a result of human factors errors vice mechanical systems. Research has shown that the number of accidents in aircraft due to mechanical or environmental problems has drastically dropped in the last several decades. For example, in 1977 a study of naval aviation revealed the proportion of accidents related to mechanical and environmental causes were practically equivalent to those linked to human error. By 1992, a follow on study showed that accidents solely caused by mechanical or environmental issues had been reduced to relatively insignificant amounts. While the second study acknowledged a decrease in both mechanical and human related error, the proportion of accidents directly attributed to failure of human factors had only dropped by 50% (Shappell & Wiegmann, 1996).

Over the past several decades, numerous technological developments in aircraft design have drastically improved safety, reducing the number of aviation related accidents. Nevertheless, human error still accounts for more than 80% of aviation mishaps (Wiegmann & Shappell, 1999). The challenge of reducing human error remains a vital concern in all high reliability organizations. Investigations into causes of accidents have revealed that the major contributors to aviation mishaps are frequently linked to some form of human error. As task complexity increases, human error becomes more prominent; especially when tasks are increasingly dependent on team interaction. Human factor causes of crashes in aviation have been commonly linked to a failure in interpersonal communication, crew coordination, decision making, situational awareness, and leadership (Flin, O'Connor, & Mearns, 2002). Considering the additional stress imposed on aviators during military flight operations, team processes and information sharing between crewmembers can become problematic. Therefore, it is important that crewmembers are trained properly for technical as well as interpersonal aptitudes.

Training in aviation has traditionally centered on technical skills and tasks. This is true for both commercial and military flight training. The philosophy remained persistent that aviators should concentrate on the essential technical skill of how to fly the aircraft. The idea of team cohesion and interaction amongst crewmembers took a proverbial backseat. The assumption was that non-technical skills required in aircrew coordination would naturally materialize without extensive and focused training. Non-technical skills are comprised of cognitive, social, and personal resource skills used to complement traditional technical skills (Flin, O'Connor, & Crichton, 2008). These competencies have proven fundamental in operating in team related tasks such as flight crew management. Non-technical skills also contribute to the safety of a mission. Fortunately, in the past two decades, crewmember teamwork and coordination have begun to emerge as priorities. This precedence started in the commercial aviation sector and quickly migrated to military flight programs.

In an attempt to combat the large number of mishaps in aviation related to human error, a training program was developed to provide training to aviation crew personnel on the non-technical skills required for safe operations. The program sought to develop competencies related to team coordination and communication in the cockpit. The training was originally called Cockpit Resource Management, but now it has adapted and is commonly referred to as Crew Resource Management (CRM). CRM was designed with the objective to reduce the number of cockpit mishaps by decreasing the number of human performance related errors (Alkov, 1989). The training has since emerged as a common method to teach interpersonal dynamics and team coordination skills (O'Connor, Hahn, & Nullmeyer, in press).

B. NEED FOR EVALUATION

Despite numerous technical advancements over the last two decades in aircraft system safety design, minimal changes have been made to the Navy’s CRM training program. Missions, personnel, and equipment have changed over the last ten years. Additionally, there is no guarantee that a training program will maintain its effectiveness over time, and thus must be evaluated frequently. In large organizations like the U.S.

Navy and U.S. Marine Corps it is challenge to establish consistency in any training program. Even with the large number of studies that previously assessed CRM effectiveness, the majority of the evaluation studies of the Naval CRM program were performed over a decade ago. Since that time, there has been progress in human factors and training research (Flin et al., 2008). Therefore, there is a need for additional research evaluating the effectiveness of the Navy's CRM program. A robust analysis helps determine if the training continues to accomplish its designed intention of improving safety and crew coordination in naval aviation. Applying methods and techniques used in previous CRM evaluation research, this thesis provided an updated evaluation of the Naval CRM program to fill important research gaps. An update review of CRM is important, as it provides insight on the vitality of the Navy's program. Furthermore, this research assessed whether the program is still positively accepted by the U.S. Naval aviation community and provided recommendations on ways the effectiveness of the training can be enhanced.

C. OBJECTIVES

The purpose of this research effort is to evaluate the effectiveness of the current Navy CRM training program across multiple aviation platforms. This in depth analysis presents valuable metrics as evidence of the practicality of the Navy's current CRM program. In addition, this study investigated differences between types of Navy and Marine Corps aviators with respect to the implementation of CRM training. The specific goals of this effort include:

- To analyze the *reactions* of naval aviators after receiving Crew Resource Management training,
- to analyze the *learning* of naval aviators receive, in terms of attitude changes and knowledge retained, as a result of CRM training
- to analyze the *impact* CRM training has on the naval aviation community in terms of improving safety.

D. RESEARCH QUESTIONS

Using a multifaceted evaluation technique in the present study allowed a number of questions regarding CRM effectiveness to be answered. Analyzing CRM on multiple levels is accomplished by investigating learner feedback as well as analyzing outcome metrics (reported safety data) across the Navy's aviation community. Specifically, this study addressed the following questions:

- What do aviators think about the Navy's CRM training program? How have the reactions changed, if at all, over the last decade?
- Are attitudes towards concepts addressed in CRM training different between junior and senior aviation personnel? Does experience affect the amount of knowledge retained from the program?
- Are there variations in attitudes or knowledge amongst aviators who fly different airframes?
- Has CRM impacted the safety mishap rate in naval aviation over the last decade?

E. METHODOLOGY

This study investigated the effectiveness of naval aviation CRM training applying a multi-level evaluation approach. Using Kirkpatrick's (1976) hierarchy of evaluation, three levels of the training assessment topology was conducted looking at the reactions, the amount of learning received and organizational results of the Naval CRM program. The learning portion is divided into two separate analyses to examine aviators' attitudes towards concepts taught in CRM and the requisite knowledge acquired from the training. Each method is briefly described below and is discussed in greater detail in the respective sections.

1. Reaction Analysis

First, reactions of naval aviators were examined using historical data provided from the Naval Safety Center. A database consisting of CRM reaction responses from fiscal years 2000 to 2008 for over 50,000 naval aviators were analyzed. The effects of rank and airframe on reaction responses were investigated to determine perception of aviators concerning the utility of CRM training.

2. Attitudinal Analysis

An attitude questionnaire based upon the Cockpit Management Attitudes Questionnaire (CMAQ; Gregorich and Wilhelm, 1993) was distributed to U.S. Navy and Marine Corps aviators attending the Naval Aviation Schools Command. The responses from the questionnaire were analyzed to examine the attitudes towards the non-technical concepts addressed in CRM training. The effect of rank and airframe on attitude responses was investigated and analyzed.

3. Knowledge Analysis

Knowledge acquired from CRM training was analyzed by developing a 10 item, multiple-choice knowledge assessment. This assessment was distributed to fleet aviators attending the Naval Postgraduate School and Naval Aviation Safety School. The score results were collected and analyzed to gauge the level of knowledge from CRM training. Again, differences in rank and airframe were considered in the CRM knowledge assessment.

4. Organizational Impact Analysis

Naval Aviation Class A mishap data, from the Naval Safety Center, was examined from 1997 until 2007 to evaluate impact of CRM as a causal factor. Mishap data also was analyzed and compared to previous studies. Comparisons were made between rotary wing and tactical jet (TACAIR) aircraft. As human error continually is identified as a causal factor in many mishaps, an examination of human error contribution was investigated to determine the role of CRM concepts in aviation mishaps.

F. HUMAN SYSTEMS INTEGRATION

In Naval aviation, system performance and flight safety are heavily reliant on successful aircrew coordination and communication. Therefore, Crew Resource

Management skills are an essential piece of HSI. Human Systems Integration (HSI) is an interdisciplinary approach that makes explicit the underlying tradeoffs in each of its domains, facilitating optimization of total system performance and reduced life cycle costs. The term Human Systems Integration is interpreted differently by various organizations. According to the Handbook of Human Systems Integration, HSI is a technical and managerial concept bringing together various disciplines with the goal of appropriately incorporating humans into the design, production, and operation of programs and systems (Booher, 2003). One of the main goals of the Navy's HSI program includes improving human performance and safety in all operational systems. The Department of the Navy currently recognizes eight key domains of Human Systems Integration including:

- Manpower – a term used to denote the number of human resources, including military and civilian, required and available to operate and maintain a system.
- Personnel – the aptitudes, experiences, and other human characteristics necessary to achieve optimal system performance. The goal is to match the "right person" to the "right job."
- Training – acquisition of the requisite knowledge, skills, and abilities needed by the available personnel to operate and maintain systems under various operational conditions.
- Human Factors Engineering – the comprehensive integration of human characteristics into system definition, design, development, and evaluation to optimize the performance of human-machine combinations.
- System Safety – the inherent ability of the system to be used, operated, and maintained without accident or injury to personnel.
- Health Hazards – the inherent conditions in the operation or use of a system that can cause death, injury, illness, disability, or reduce job performance of personnel.

- Human Survivability –a system characteristic that allows personnel to exist and function during and following exposure to hostile environments or situations. Includes issues involving enemy and friendly combat weapons-induced injuries and the inherent hazards to personnel during threat/combat conditions, and the inherent hazards of military equipment
- Habitability – the physical environment in which personnel are required to live, work, and sleep while performing their duties. This includes the physical and psychological needs of the individual and group, and takes into account morale and the social environment during both sustained and continuous military operations.

An evaluation of the Navy's CRM training program is specifically relevant to five HSI domains: Manpower, Personnel, Training, Human Factors, and System Safety. With respect to the domains of Manpower and Personnel, this study provided information on differences in the naval aviation workforce in terms of their acceptance towards CRM training. In addition, the knowledge, skills, and attitudes acquired as a result of the Naval CRM program were assessed. The outcomes of this study provide insight into selection criteria for crew team members across aviation platforms. This may help address future workforce challenges and speaks to Manpower, Personnel, Training, and Education (MPTE) strategies for competency management. Additionally, this evaluation helps identify how CRM training might be implemented elsewhere in the Navy organization to teach aspects of human factors during operations.

The HSI domains of human factors and system safety also are relevant to the study described in this paper. Analyzing CRM effectiveness has implications for aircraft system design. Just as mechanical improvements have reduced system errors and increased system safety, the CRM program aims to minimize the potential for human errors. Specifically, it seeks to reduce the crew coordination failures that lead to accidents and injuries. These errors can result from human to human interaction as well as human to machine misinterpretations. Another benefit of CRM training is that it provides a risk management and assessment tool for aviators and crew personnel. CRM has a direct impact on a team decision-making process. Optimizing risk management will ensure system safety and limit preventable errors. Finally, examining mishap rates in aviation over the past decade may reveal undiscovered trends in flight safety data directly accredited to the impact of CRM training.

G. THESIS ORGANIZATION

A brief literature review of the history and development of CRM training is discussed in Chapter II. This review encompasses a breakdown of the seven non-technical skills taught in the Naval Crew Resource Management aviation program. Historical methods used to evaluate training effectiveness also are discussed in Chapter II. This thesis evaluated the effectiveness of the Naval CRM program using a multifaceted approach. The subsequent chapters are devoted to analyzing the CRM program on multiple levels. Chapter III describes an analysis of the reactions of trainees to the CRM program. Chapters IV and V outline a method to gauge the level of learning received from Naval CRM. These two chapters are broken down into two respective learning assessments: attitudes and knowledge. Chapter VI is devoted to looking at the organizational impact CRM has had on naval aviation over the past decade. This included an analysis of historical naval aviation safety mishap data. Chapter VII summarizes the results from each of the four distinct analyses and discusses future implications for the U.S. Navy. Additionally, a few recommendations are made for further research in evaluating and applying CRM concepts in other communities.

II. LITERATURE REVIEW

A. HISTORY OF CREW RESOURCE MANAGEMENT

Concepts included in Crew Resource Management (CRM) can be traced back to human factors research conducted in military aviation during World War II. During this period, highly reliable aircraft were being introduced and more failures were being acknowledged as operator oversight rather than mechanical defects (Davies, 2001). This recognition of crew error prompted extensive research directed at identifying various types of performance skills, unrelated to mechanical error, to ensure pilot and aircraft safety (NASA; Cooper, White, & Lauber, 1980). While CRM concepts and research may have started in the military, it was the civilian and commercial aviation industry that began to address the 'human' aspects of aviation. The CRM training program was first instituted and developed for civil air carriers.

The definition of Crew Resource Management has taken several forms since its origination. Although a universal definition has not been reached, many agree that development of teamwork attitudes, personalities, and social interactions are at the center of CRM training. Others foresee CRM training as a method to foster proper crew behavioral skills (Wiener, Kanki, & Helreich, 1993). The term "behavioral skills" falls subject to a wide range of interpretation from program to program and, therefore, contributes to lack of agreement on what really defines CRM (Salas, Rhodenizer, & Bowers, 2000). It is easy to focus on the 'crew' aspect of CRM, but the 'resource' is equally imperative to the program's success. In fact, the Federal Aviation Administration (FAA) recognized that CRM training and applications must include all resources to achieve effective coordination. This includes people, equipment, cognitive processes, procedures and information. The definition offered by the FAA (1989) states:

> CRM can be broadly defined as the utilization of all available human, informational, and equipment resources toward the goal of safe and efficient flight. CRM is an active process by crewmembers to identify significant threats, to communicate them, and to develop, communicate, and carry out a plan and actions to avoid or mitigate each threat. CRM also deals directly with the avoidance of human errors and the

> management and mitigation of those errors that occur. CRM reflects the application of human factor knowledge to the special case of flight crews and their interactions with each other, with other groups and with the technology in the system (p. 2).

Many still agree that CRM is a technique that ensures the proper utilization of all available resources to effectively achieve a desired outcome. Salas et al. (1999) further expound on the definition of CRM by describing it as “a set of instructional strategies designed to improve teamwork in the cockpit by applying well-tested tools (e.g., performance measurements, exercises, feedback mechanisms) and appropriate training methods (e.g., simulators, lectures, videos) targeted at specific content (i.e., teamwork knowledge, skills, and attitudes).” These founding principles were the catalyst as CRM became quickly adopted throughout civilian and military aviation, including the Navy.

1. Civilian and Commercial Origins

CRM training programs originated in the United States starting with a workshop in 1979 sponsored by the National Aeronautics and Space Administration (NASA; Cooper et al., 1980). There was widespread trepidation concerning the large number of commercial aviation accidents attributed to “pilot error” (Helmreich, Merritt, & Wilhelm, 1999). Failures in decision making, interpersonal communications and leadership were identified as some of the most common human errors linked to accidents in aviation. The objective of developing a resource management program was to train aviators on aspects of interpersonal communications, decision making, and leadership. The program quickly spread throughout many civilian aviation applications across the globe (Helmreich, et al., 1999). At the program’s onset, it was referred to as Cockpit Resource Management. The term “Cockpit” was largely due to the fact that sources of human error develop from failures in coordination between pilots and their team members in the cockpit. The commercial airlines led the effort to establish the resource management program for a number of reasons. The recognition of a need was a result of the research conducted at NASA looking at training aircrew members in skills beyond their technical flying aptitude (Lauber, 1987). Another basis for CRM training came from public concern over the increased frequency of fatal aircraft accidents in the 1970’s due to human error (Prince & Salas, 1993).

The first generation of Cockpit Resource Management training started with a United Airlines program in 1981 (Helmreich, et al., 1999). This training consisted of seminar-type forums looking at managerial issues and leadership concerns. Much of the attention was focused on problems dealing with the lack of assertiveness in subordinates and the potential dictatorial behavior of commercial airline captains (both of which have been identified as potential causal factors in crashes). Many of these modules highlighted the need for recurrent training, even for experienced pilots. In addition, simulators were beginning to be utilized to observe and practice CRM techniques.

In 1986, NASA convened once again to discuss expanding the Cockpit Resource Management Program (Orlady & Foushee, 1987). This conference brought together many organizations throughout the world to discuss implementation of CRM type courses. At this time, the application of training began emphasizing the need to integrate more teamwork and coordination concepts. This transformed the program from being referred to as “Cockpit” to “Crew” Resource Management. Additionally, the civilian airline industry had recognized the modular nature of aviation crews (Helmreich et al., 1999). Flight crews were being assembled on an ad hoc basis with little continuity. Recommended changes to the program included incorporation of flight attendants, dispatchers, and maintenance personnel. This set in motion joint cockpit and cabin CRM training. Also, checklists were being added by many of the airlines that included behavioral skills. These checklists re-emphasized the importance of teaching non-technical skills to improve flight crew coordination and safety.

It was not until the late 1980s and early 1990s that the commercial aviation industry began upgrading CRM training to include recurrent classroom training in conjunction with performance assessments (Helmreich & Foushee, 1993). Shortly thereafter, CRM skills were starting to merge with existing technical instruction. Advanced training quickly emerged for check airmen and instructors with regards to CRM concepts (Helmreich et al., 1999). With so many different CRM related safety programs surfacing in the commercial aviation sector, the FAA launched the Advanced Qualification Program (AQP). The FAA describes AQP as a "systematic methodology for developing the content of training programs for air carrier crewmembers and dispatchers" (Federal Aviation Administration, 2004). AQP is designed to focus flight training and evaluation on crew as well as individual performance. The goal of AQP was to decrease the likelihood of crew-related errors by encouraging innovative ways to improve modify known causes of human error. Further, it eradicates some previous regulatory barriers integrating CRM and technical training, and provides incentives to organizations to adopt collective programs (Helmreich & Foushee, 1993).

Helmreich et al. (1999) estimated CRM had evolved into its fifth generation by the late 1990s. These subsequent generations are realizing that human error is inevitable. Because errors are certain, the more recent notion is that CRM can be implemented as a means to not only minimize the occurrences of human error, but incorporate a technique to identify error early and mitigate consequences of those errors (Flin et al., 2008). More recently, CRM is being reviewed and updated to ensure it is properly incorporated with today's technology. Simulation and computer based programs both provide a unique medium for future CRM training. CRM concepts and applications are the most well known in commercial and military aviation; however, CRM concepts and programs continue to expand into all types of industries. Today CRM techniques are being applied and commonly used throughout many high reliability industries such as oil production platforms, nuclear energy, and numerous medical fields (see Flin, O'Connor, & Mearns, 2002 for a review).

2. Military Applications

In the early days of CRM training, the military believed it to be a civilian-specific enterprise. Leaders in the military may have initially overlooked the need for a crew training program like CRM because they already had methods in place to address coordination in the cockpit (Brown, 1987). These ranged from aircrew coordination checklists to briefings and lectures, but no formal training existed. As human error awareness grew in the civilian aviation community, the military's view of human error rapidly changed. As with commercial aviation, research of military aviation mishaps showed that human-related failures were being identified as the main causal factor (Leedom, 1990). In addition to saving lives and retraining personnel, avoiding accidents helps save millions of dollars in aircraft replacement. The U.S. military services started their own version of CRM training in the late 1980s and early 1990s and these are discussed briefly below.

United States Air Force. The U.S. Air Force was one of the first services to implement its own form of CRM training in the mid 1980's within its Military Airlift Command, now known as Air Mobility Command (MAC). The Air Force Inspection and Safety Center subsequently developed a CRM type course in 1990 for all single seat fighters (O'Connor, Hahn, & Nullmeyer, in press). The course was originally labeled "Aircrew Coordination Training," or ACT. Currently, CRM training is mandated for all Air Force crewmembers. In the Air Force Instruction (AFI) 11-290, it states the goals of the Air Force's CRM program are to maximize operational effectiveness and combat capability while preserving personnel and material resources (Secretary of the Air Force, 1995).

United States Army. The Army introduced the first service-wide, standardized CRM training program in the early 1990s, also known as ACT. The U.S. Army defines the aircrew coordination as "a set of principles, attitudes, procedures, and techniques that transforms individuals into an effective crew" (Katz & Grubb, 2003). The U.S. Army Aviation Center (USAAVNC), the Army Research Institute (ARI), and the U.S. Army Safety Center (USASC) conducted a number of studies researching crew-level performance to develop a training program that would enhance team aircrew coordination. Since that time, the Army has updated its program by implementing the ACT-E (Aircrew Coordination Training – Enhanced) program in 2006. Currently, all army aviators are given initial ACT-E training with annual refreshers (O'Connor, Hahn, & Nullmeyer, in press).

United States Navy and Marine Corps. The U.S. Navy and Marine Corps also demonstrated interest in improving aircrew coordination and began developing its own Crew Resource Management program in the late 1980s and early 1990s. In 1991, the

Navy initiated a research and development effort at the Naval Air Warfare Center Training Systems Division (NAWCTSD) to identify common behavioral skills associated with human error mishaps in aviation. Drawing on prior success in civilian programs, seven skills were identified and integrated into the initial form of CRM also known as the Aircrew Coordination Training (ACT) program. Realizing the need to train its aviation workforce in non-technical coordination skills, the CNO mandated CRM training in 1995 for all Navy and Marine Corps aviation personnel (OPNAVINST 1542.7A). The training consists of a two to three day initial course using lectures, practical exercises, videos and case studies. To this day, every crewmember in naval aviation is required to receive the introductory CRM course as well as annual training refreshers. Naval aviators also must have their CRM skills evaluated by a certified CRM instructor. The Navy's CRM program is discussed in more detail below.

B. CRM IN NAVAL TRAINING

1. Naval CRM Goals

Safety was the primary reason behind the development of the Naval CRM training program (Prince & Salas, 1993). The goal of the Naval CRM training program is to provide aviators with the necessary skills in order to "improve mission effectiveness by minimizing crew preventable errors, maximizing crew coordination, and optimizing risk management" (Chief of Naval Operations, 2001). Since human error still accounts for approximately 80% of all naval aviation mishaps (Naval Safety Center, 2006), the mission of naval CRM training is to reduce the number of human related errors in aviation. There are a multitude of methods and ways to approach changing attitudes and behavior in an organization. The Navy sought to develop an effective academic approach to educate and train its personnel in CRM non-technical skills. CRM concepts and techniques would be taught in classrooms and evaluated through the use of simulators and instructional flight training. In order to achieve this objective, the Navy desired a standardized training strategy to incorporate basic principles of CRM training throughout naval aviation.

2. Evolution of CRM in Naval Aviation

There are over thirty different types of aircraft used in Navy and Marine Corps aviation, ranging from single seat jet fighters to multi-crew rotary airframes. The number of different airframes makes applying CRM concepts more challenging than commercial aviation. Designing consistent and standardized naval CRM programs is problematic. When the Navy began mandating CRM training, it appeared to have been accepted with more enthusiasm in some aviation communities as compared to others (Prince & Salas, 1993). This disparity may have been due to the perceived relevance of the training from one community to another. The observed lack of enthusiasm for CRM training in some airframes launched vast amounts of research in evaluating potential methods for tailoring CRM programs to specific communities. Many studies focused on the manner in which programs could be designed and implemented to account for the diversity within naval aviation (Oser, Salas, Merket, & Bowers, 2001).

Identifying fundamental behavioral skills in naval aviation was the initial step taken to develop a successful CRM training program. Oser et al. (2001) examined existing CRM training programs, evaluated naval safety mishap reports, and conducted interviews with subject matter experts (SMEs) to pinpoint some of the critical skills involved with aviation performance. Essentially, their research consisted of a needs analysis aimed at identifying what competencies were required across a number of aviation communities. In addition, several prototype CRM training programs were developed to further assess the necessary skills for effective crew coordination and team performance in the cockpit (Fowlkes, Lane, Salas, Franz, & Osler, 1994). Together, these studies led to the identification and evaluation of seven skill areas required for effective aircrew coordination (Prince & Salas, 1993). The seven skills are still currently used in Naval CRM programs and discussed in further detail in the following section.

3. Seven Critical Skills of Naval Crew Resource Management

Aviation tasks routinely involve skills beyond traditional technical training. Through research and practice, certain non-technical skills were identified as critical to effective crew coordination in naval aviation. Many of these skills were documented using critical incident reviews with over 200 naval aviation crewmembers to identify required team processes (Prince & Salas, 1989). From these studies, seven skills were acknowledged by naval aviators as significant proficiencies used during flight operations. These skills included decision making, assertiveness, mission analysis, communication, leadership, adaptability/flexibility, and situational awareness. This section provides a brief description of each CRM skill currently taught in the Navy's Crew Resource Management Program.

a. Decision Making

Decision Making is defined as the skill of selecting a particular course of action using logical and sound judgment based upon the available information (Alkov, 2005). An important component in deciding and executing a specific course of action comes from understanding, interpreting, and validating gathered information. This includes aspects of assessing and verifying the problem, identifying potential solutions, anticipating consequences, informing others of the decision and rationale behind it. Effective decision-making involves transforming provided information or data into planning and action. Variables required in the decision-making process include, but are not limited to: level of uncertainty, experience, time, risk assessment and management.

Safety is critical in high-risk organizations and because decision making can be influenced by time and stress, good decision making skills are important in these domains. Aviation is no exception. A study by the National Transportation and Safety Board (NTSB) in 1991 revealed that 47% of aircraft accidents in the United States between 1983 and 1987 involved poor crew judgment and decision-making (from Flin et al., 2008). There have been numerous failures in aviation attributed to a lack of timely and effective decision making. The result has expounded research in psychology about the types of decision making depending on the environmental conditions. The classic rational decision models, where adequate time was available to consider an optimal solution, lacked applicability to high stress situations. Many psychologists began looking at decision making in time-pressured, demanding, and risky situations. The outcome included the theory of Naturalistic Decision Making (NDM) (Klein, Orasanu, Calderwood, & Zsambok, 1993). NDM research is dedicated to choosing a course of action under high ambiguity conditions, or when there is extensive pressure and little time (Salas & Klein, 2001). The NDM process is relevant when analyzing decision making in naval aviation as situations arise requiring quick and accurate responses.

Teamwork including crew coordination and communication are considered key factors of good decision making applied to team-related tasks. Methods to improve team decision-making skills include system or organizational design and formal team training (Orasanu & Salas, 1993). These decision-making skills are put to the test through simulated team exercises and evaluating communications which develop team mental models. Naval CRM training incorporates different decision-making approaches and techniques to effectively manage resources, workload, and information within military aircrews (Flin et al., 2008).

b. Assertiveness

Assertiveness can be classified as a willingness and readiness to actively participate in a specific task (Naval Aviation Schools Command, 2008). In aviation, this involves the ability to state and maintain your position (or point of view) until convinced by facts, rather than the authority of another individual, that another viable option is preferred. Assertiveness entails expressing one's feelings, concerns, or ideas with a clear and direct approach (Jentsch & Smith-Jentsch, 2001). Aircraft accidents have been attributed to lack of assertiveness, where voice data recorder tapes indicated a co-pilot's reluctance to question a decision made by the captain (Flin et al., 2008, p. 79). In the military, barriers preventing proper assertiveness can include rank, gender, experience, or fear of reprisal.

Assertiveness is applied as both an attitude and a skill, requiring both initiative and courage to act (Jentsch & Smith-Jentsch, 2001). It is important for assertive qualities to be distinguished from passive or aggressive methods for resolving conflict. Passive refers a failure to stand up for one's opinion or beliefs. Often passivity is a practice used when trying to avoid conflict or confrontation. Frequently, passive attitudes and behaviors do not work very well in high-risk situations. Aggressiveness, on the other hand results in a disregard for other’s opinions and input. The behavior can be exhibited through both non-verbal and verbal actions. A significant distinction between assertive and aggressive behavior is that aggressiveness can be overly dominant, intimidating and hostile. However, aggressive attitudes are required in some circumstances such as emergencies (Flin et al., 2008). CRM training seeks to educate aviators on recognizing the differences between passive, assertive, and aggressive and what situations are appropriate for each. Moreover, assertiveness training in CRM programs can be evaluated through use of behavioral role modeling and simulation (Foushee & Helmreich, 1988; Jentsch & Smith-Jentsch, 2001).

c. Mission Analysis

Mission analysis is centered on the ability to make short-term, long-term, and contingency plans (Naval Aviation Schools Command, 2008). Teaching the concept of mission analysis in CRM focuses on ways to understand how to coordinate, allocate and monitor crew and available aircraft resources. Aviation mission analysis is commonly conducted in the three phases of flight: pre-flight, in-flight, and post-flight. Pre-flight analysis involves initial planning and briefing in preparation for the mission. CRM concepts touch on the need for adequate consideration of foreseeable challenges and risks associated with the mission. In-flight analysis deals with monitoring the progress of the mission and identifying any necessary changes or alterations as situations arise. CRM highlights that successful in-flight analysis uses the knowledge and capabilities of the entire crew and the resources at their disposal. Post-flight addresses the proper review of the operation using timely and interactive debriefs. This provides a way to address both crew and individual lessons learned from the mission.

While mission analysis is important in each stage of aircrew flight operations, the pre-flight planning, prioritizing, and assigning of tasks is perhaps the most crucial stage (Prince & Salas, 1993). The pre-mission brief will set the course for the events to follow. Orasanu (1990) conducted an experiment with commercial airline flight crews and found that crews with supplementary emergency plans and contingency strategies ultimately were more effective. Mission analysis remains a valuable skill taught in naval CRM to enhance organization and planning in order to reduce error in aircrew coordination.

d. Communication

Communication is the clear and accurate sending and receiving of information, instructions, or commands, and providing useful feedback (Naval Aviation Schools Command, 2008). Communication is considered a non-technical skill reflective of an organization's culture and policy, and enhanced through training (Flin et al., 2008). The types of communication within a flight crew can be verbal, non-verbal, gestures, written, or simple voice inflections. Further, communication can be affected by age, gender, rank, culture, attitude, and task or workload. For example, in naval aviation, landing aircraft on naval aircraft carriers is a very complex and dynamic operation. Outside the range of external environmental factors, effective communication between the ship's bridge, landing deck and aircraft crew is essential to success. Words, tones, and gestures play a significant role in the operation. Research has shown that a breakdown in communication in aircrews has been linked to causal factors in many Navy and Marines Corps mishaps (Hartel, Smith, & Prince, 1991).

Effective communication is comprised of four basic elements: clarity, accuracy, timeliness, and usefulness. Clarity, or what Flin et al. (2008) describe as explicitness, is clearly stating a desired or intended action while minimizing ambiguity. Operational communication requires efficiency; not overloading the receiver with too much extraneous information but providing enough to plainly express what information is critical for achieving the mission. This is known is naval aviation as 'comm-brevity'. Accuracy is making sure the information transmitted from the sender to the receiver is correct. The idea of thinking about, or verifying, the facts before you transmit information enhances clear communication. Errors occur when the wrong information is passed from one source to another individual or a chain of team members. The third basic element is timeliness. Timeliness does not imply that all information should be disseminated as quickly as possible. In fact, effective timing knows when information should be sent to the receiver. This largely depends on the task and workload of the individuals involved (Flin et al., 2008). Studies have shown that high performing aircrew communicated verbally less often under high workload situations as compared to normal flight operations (Orasanu, 1993). Effective flight briefs underscore the essential and relevant information pertinent to the mission. Communication establishes the foundation for the other CRM skills and is an important component leading to successful mission accomplishment in aircrew coordination.

e. Leadership

Leadership is taught in naval aviation as the ability to direct and coordinate the activities of crewmembers; encouraging the crewmembers to work together as a team (Naval Aviation Schools Command, 2008). A delicate balance of effective leadership experience, combined with valuable management skills, is an integral part of driving any team to success. Naval CRM identifies two distinct types of leadership styles: designated and functional. A designated leader is someone who possesses the responsibility to make final decisions. In contrast, a functional leader arises when the time for leadership is needed or required. Functional leadership is typically a result of prior experience and knowledge. Generally, teams have a single leader. However, oftentimes leadership skills expand beyond the designated leader to the other team members. Therefore, leadership is an important skill for all members of teams.

A large component of good leadership in aircrew coordination is the ability to balance the workload with the resources of a crew (Flin et al., 2008). In military teams, leaders take on increased responsibility to ensure the team achieves the mission in a safe and effective manner. A study carried out by Thordsen et al. (1990) suggested that micromanagement of Army helicopter aircrews led to poorer performance than those teams that were not micromanaged when evaluated in simulated missions. One common method to assess leadership skills in aviation is through the observational ratings. Ratings of aviators are carried out using behavioral markers that include a section on leadership and management (see Table 1).

Table 1. Leadership Skill Evaluation Using NOTECHS Behavioral Markers (from Flin et al., 2008, p. 151)

Category	Elements	Example behaviors
Leadership and managerial skills	Use of authority and assertiveness	Takes initiative to ensure involvement and task completion
	Maintaining standards	Intervenes if task completion deviates from standards
	Planning and co-ordinating	Clearly states intentions and goals
	Workload management	Allocates enough time to complete tasks

f. Adaptability/Flexibility

Adaptability and/or flexibility describe the ability to alter a course of action when new information becomes available (Naval Aviation Schools Command, 2008). Many situations in aviation arise where a deviation of a particular task is warranted. Emergencies are a prime example of the need for adaptability. CRM focuses on ensuring that crewmembers are aware of planned missions and understand the possibility of change. Offering alternative solutions to unexpected circumstances may be required. Aircrews must be flexible to change anticipated courses of action in order to respond to demands of a particular mission or task. Research has indicated that lack of flexibility and adaptability in naval aviation has contributed to Navy and Marine Corps mishaps (Hartel et al., 1991). Adaptability also is related to leadership in dynamic environments like naval aviation. As situations change, in terms of unexpected risk, weather, or degraded communications, leadership skills must adapt to meet those challenges. CRM instruction discusses ways to increase adaptability, including more effective mission planning analysis, preparing for multi-role missions, and expecting some variations in execution.

g. Situational Awareness

A common definition of situational awareness stems from Endsley (1995) as she describes it as “the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future” (p. 36). In Naval CRM programs, situational awareness is the degree of accuracy by which one's perception of the current environment mirrors reality (Naval

Aviation Schools Command, 2008). It is a cognitive skill based in perception and attention. Situational awareness can be influenced by a myriad of individual factors including processing limitations, memory, and mental models. Research has shown that in a large number of commercial aviation accidents attributed to human error, problems with situational awareness played a significant role (Endsley, 1995).

Situational awareness as a skill encompasses being conscious and understanding of what is happening both inside and outside the aircraft. Although its roots derive from air combat maneuvering and military battlefields, situational awareness has been adopted as a means of improving safety in a variety of vocations. Effective situational awareness means staying oriented in time and space, focused on requirements. Human factors that can degrade situational awareness include: insufficient communication, fatigue, stress, task load, groupthink, or degraded operating conditions. Cognitive skills are difficult to train and even more difficult to assess. The Naval CRM program addresses situational awareness through methods the brain uses to store, process, and recall information.

The seven skills are the core elements used in the Navy to train its aviator crew personnel the competencies for successful crew resource management. Because these fundamental skills are non-technical in nature, it remains challenging to measure just how lucrative a CRM training program can be; especially with a course based heavily in cognitive and behavioral proficiencies. In order to productively evaluate the effectiveness of the CRM program, a logical process must be applied. The next section discusses some common methods and approaches to gauge utility of CRM training in order to decompose the evaluation into reasonable steps.

C. EVALUATING CREW RESOURCE MANAGEMENT

1. Purpose of a Training Assessment

Evaluating a training program can indicate whether the training is meeting its intention and goals. Analysis may include a review of the program’s objectives, methods, feedback, and results in terms of its overall effectiveness. Goldstein (2002) defines training evaluation as a “systematic collection of descriptive and judgmental information necessary to make effective training decisions related to the selection, adoption, value and modification of various instructional activities" (p. 138). Training is futile if it does not make a difference that provides practical relevance and value to an organization (Holt, Boehm-Davis, & Beaubien, 2001). The impact of training may be seen on one or more distinctive levels including the individual, the team, and the organization. As stated earlier, the purpose and goal of CRM training is to improve safety of a crew through concepts directed at improving coordination, communication and awareness. An effective evaluation helps validate and determine if the current CRM program positively impacts human capital and maximizes use of available resources in naval aviation.

2. Common Evaluation Techniques

There are several techniques used to evaluate the effectiveness of a training system or program. Most of these methods address issues of learning and retention, measuring on-the-job performance, and investigating the value added in terms of cost benefit. The objective of the training being conducted provides a fundamental basis for selecting an appropriate evaluation process. Specific methods used to evaluate programs are largely dependent on the knowledge, skills, or attitudes outcomes desired as a result of the training. Consideration must be given to whether the training is primarily directed at task-related or team-related competencies. Most task-related training relies on Transfer of Training (ToT) evaluation techniques, measuring precise performance criteria (Roscoe & Williges, 1975). However, CRM is predominantly based on instructing behavioral skills and therefore requires a multifaceted evaluation approach.

A multifaceted evaluation approach is preferred when examining CRM training effects in aviation as it considers several levels of assessment (Gregorich & Wilhelm, 1993; Kraiger, Ford, & Salas, 1993). The major focus of this thesis is founded on evaluating CRM's validity and its effects on participants' reactions, effect of learning, and organizational outcomes. In order to examine the effects of CRM on naval aviation, it is essential to distinguish between what impact the CRM program has on aviators in terms of their reactions, attitudes and knowledge. The present study considers how CRM training continues to affect the Navy as an organization in terms of improved safety and reduction of mishaps.

3. Previous Navy CRM Assessments

As with the commercial aviation industry, there has been a multitude of literature written, and studies evaluating, aspects of CRM and its effectiveness in naval aviation (see O’Connor, Campbell, Newon, Melton, Salas, Wilson, 2008; Salas, Burke, Bowers, Wilson, 2001; Salas, Wilson, Burke, & Wightman 2006; for reviews). These studies looked at the usefulness of training on distinctive levels by assessing the impact of training ranging from individual to organizational training effects (O'Connor, Hahn, & Salas, in press). Only a few of the studies applied a multi-faceted approach to their evaluations. However, many of these studies analyzed CRM effectiveness in term of participants’ reactions to the training, their changes in attitudes and behaviors, as well as the impact the training had on the organization. O’Connor, Hahn, & Salas (in press) summarized eight studies that previously analyzed naval aviation CRM training (see Table 2).

Table 2. Prior Studies of CRM Training Effectiveness in Naval Aviation (From O'Connor et al., in press, p. 5-6).

Author	Participants	Reactions	Learning	Behaviour	Organisation
Alkov (1991); Alkov & Gaynor (1991); Alkov (1989)	90 aircrew		Positive shift in attitudes.		Aircrew error mishap rate declined.
Baker et al. (1991)	41 helicopter pilots.	Positive reaction			
Brannick et al (1995)	51 aircrew			Better than average in performing CRM behaviors	
O'Connor (under review)	371 naval aviators, and 272 Navy divers.		Aviators significantly more positive attitudes to CRM than divers.		
Salas et al. (1999)	35 pilots & 34 enlisted helicopter aircrew	Positive reaction.	No significant difference in attitudes pre- and post-training. Significant increase in knowledge of CRM principles.	Trained crews performed better than untrained crews as measure by TARGETs.	
	27 helicopter pilots (12 serving as controls).	Positive reaction	No significant difference in attitudes between trained and controls. Trained aviators scored significantly better on knowledge test than controls.	Trained better during pre-flight brief, and greater number of teamwork behaviours during high workload as measured by TARGETs.	
Stout, Salas & Folkes (1997)	42 student aviators (20 experimental, 22 control)	Positive reaction	Positive shift in attitudes. Significant increase in knowledge of CRM principles	Trained crews performed better than untrained crews as measure by TARGETs.	
Stout, Salas & Kraiger (1996);	12 helicopter pilots (10 serving as a control group) .	Positive reaction	Positive change in attitudes, but not significant. No significant difference in scores on knowledge test than controls.	Trained participants performed on average 8% more desired behaviours than control as measured by TARGETs.	
Wiegmann & Shappell (1999);	290 naval aviation mishap (1990-96) causal factors				56% of the mishaps had at least one CRM causal factors. Comparable to the 58% air crew error rate found by Yacavone (1993) in an examination of 308 naval aviation mishaps (1986-90).

The CRM studies (listed in Table 2) assessed, at a minimum, one level of analysis corresponding to Kirkpatrick's (1976) hierarchy of training evaluation. This hierarchy is a commonly used method to assess the effectiveness of a training program across multiple levels within an organization and is discussed in further detail in the following section.

4. Applying Kirkpatrick's Topology of Evaluation

Kirkpatrick's (1976) topology centers on four fundamental levels of assessment. These include measuring and analyzing participants' reactions, learning, behavior, and results or outcomes (see Figure 1). Since CRM programs focus on non-technical behavioral skills, Kirkpatrick's approach has been chosen as a useful framework for evaluating naval aviation CRM training at multiple levels.

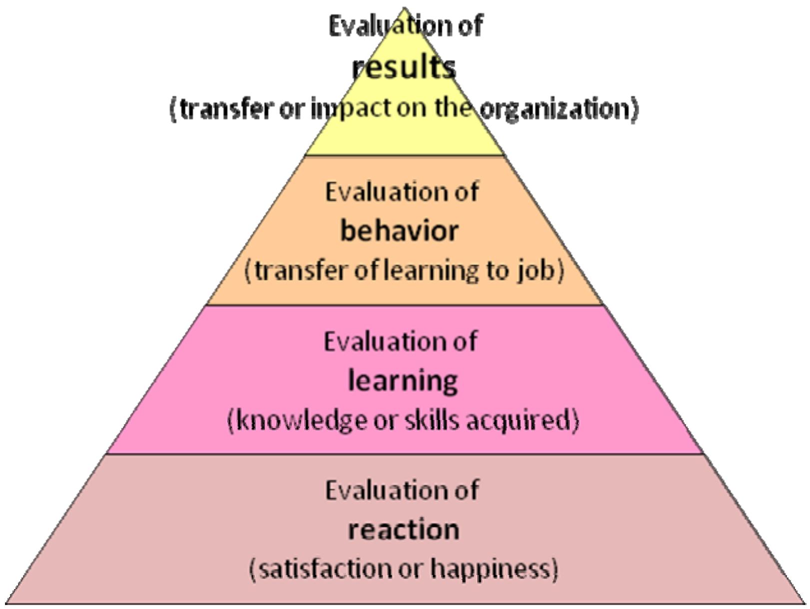

Figure 1. Kirkpatrick's Hierarchy of Evaluation

a. Reactions

Kirkpatrick's first level of training evaluation, "reactions," is concerned with simply the satisfaction of the participants with the training. It is a measure of whether the reactions of the trainees were positive or negative. In essence, it is a way to determine customer satisfaction and often referred to as a "happy sheet." Conventional methods of quantifying reactions are typically performed through the use of surveys. Survey questions are centered on assessing whether the participants enjoyed a particular training program. Responses can be indicated as simple as marking one single check on a rating scale. For example, 'how did you like the training?' Another means for determining trainee reactions is through use of a detailed inquiry into an assortment of opinions, such as open-ended responses to specifics about the training. For example, “what part of the training was of most value to you?” However, most reaction data for CRM programs are collected and evaluated using a simple Likert scale response survey or through asking participants to rank order perceived utility of CRM components (Salas, Burke et al., 2001).

Kirkpatrick (1996) contends that assessing how people feel about the training is important for two primary reasons. First, management decisions regarding training are often based on feedback provided from trainees (Kirkpatrick, 1996). If there exists a general feeling that the training seems useless or unpleasant, it may be subject to termination. Managers and supervisors deeply desire for their employees to enjoy the training in their organization. Besides, training is not cheap. A great deal of resources, manpower and money, are directed towards successful training programs. Organizations genuinely yearn for these investments to be worthwhile so as to avoid wasting time and assets on training, potentially leading to pessimistic attitudes.

Secondly, it is important to understand trainees’ reactions simply because reactions may influence learning. Research has been unsuccessful in directly linking satisfaction with the effectiveness of a training program (Mathieu, Tannenbaum, & Salas, 1992; Alliger & Janak, 1989). However, few would argue that it makes sense to conduct a training program with unfavorable reactions. Positive responses, unlike negative, are rarely harmful to the learning process. In fact, many would argue that positive reactions may enhance an individual’s motivation to learn (Salas, Burke et al., 2001). An individual who enjoys a training course can become readily enthusiastic and receptive to learning. On the other hand, if a person feels the training is of no value, it can reduce the likelihood learning took place (Kirkpatrick, 1996). It is important to realize negative reactions are not always adverse. Honest reactions can inform decision makers for potential areas of improvement in a course or prospective changes to instruction methods.

Assessing the reactions of participants is common practice in many training organizations. Unfortunately, analyzing reactions is often misconstrued as a method to evaluate a program in its entirety. Just because a training program receives enthusiastic reviews does not correlate to the value added or make the training necessarily effective (Kirkpatrick, 1996). Likewise, learning and change can occur even if the participant does not particularly enjoy the instruction. Organizations commonly fail to conduct necessary higher-level evaluations (Lovell, 2007). Reaction assessment, although important and constructive, is only the first step in the evaluation process, and further efforts are needed to assure learning and outcomes are going to result from the training.

b. Learning

Learning is the second tier of evaluation in Kirkpatrick's hierarchy. There are two distinct characteristics of measuring learning: knowledge (or skill) and attitudes. First, learning can be expressed as the degree to which knowledge or skills were acquired as a result of the training (Kirkpatrick, 1996). A training program can affect learning through cultivating both declarative and procedural types of knowledge. Declarative knowledge refers to the factual information that can be encoded and recalled from one's memory. Procedural knowledge addresses applying a specific learned behavior to a situation or task (Holt et al., 2001). The most common method to assess the knowledge and skills gained from a training program is to measure performance, typically via a test or examination. Standardized tests are popular and efficient, but do not allow for flexibility from program to program. Organizations are unique and therefore require specially designed assessment tools.

Next, learning deals with changes in attitudes. The objective of a training program may be to change attitudes towards policies, procedures, the organization, or the job itself (Phillips, 1991). If the training is designed for attitudinal modifications, an evaluation of these attitudes is warranted. The effect of time on individuals' attitudes can also play a factor. Therefore, measuring trainee's attitudes can be tricky and complex. Attitudes can be measured using surveys, focus groups, interviews, and through observation. The purpose of attitude assessments can provide the organization insight on selecting and developing personnel, ways to improve a training program, gauge progress and monitor trends.

Measuring the amount of learning transfer is much more complex than measuring reactions. It is important to note that learning does not take into account the transfer of knowledge, skills, and attitudes (KSAs) to a specific job, rather how well these attributes were acquired and retained. The application of KSAs is more relevant when addressing the behavioral outcomes of the trainee and is included in the following discussion.

c. Behaviors

The third level in Kirkpatrick's hierarchy, behavior, deals with the degree to which the training changed or altered the performance associated with a job or skill. A widely used technique to evaluate the behavior of trainees is carried out through the use of a structured observational system (Flin et al., 2008). Observational data can be a useful and important metric of performance, but the data can be subjective and may not be an accurate measure of how the participants would behave in the 'real world' when they are no longer being observed. Although helpful, behavior evaluation methods are tedious and lengthy. Monitoring job performance in aviation presents its own challenges. It requires a large amount of manpower and is resource intensive. For these reasons, a behavior analysis of CRM is omitted from this research effort, yet recommended for consideration in future efforts.

d. Results

The highest, and most powerful, measure of training effectiveness is to examine impacts at the organizational level (Kirkpatrick, 1996). Kirkpatrick uses evaluation of results to assess the outcomes of a program and whether the training concepts had any bearing on the overall objectives. Results are linked to direct impact the training made on the organization as a whole. Some outcomes are relatively easy to measure. For example measuring quantitative metrics such as production output or profit margin for a corporation is fairly straightforward. Others can be more complicated to assess, like improved customer satisfaction and employee morale. Perhaps the most popular ways to measure training outcome or results is through cost savings or return on investment (ROI) studies. ROI attempts to assess the value added to an organization.

However, in naval aviation the purpose of CRM training is to improve mission effectiveness by reducing possible errors. The cost associated with replacing trained aircrew not to mention the multi-million dollars in aircraft makes CRM worthwhile if a single accident is avoided.

Measuring mission effectiveness is not easy to assess as there are many variables involved. O'Connor et al. (2002) looked at 48 separate CRM studies in commercial and military aviation. Many of these studies evaluated CRM on multiple levels. Yet, only 17% of the studies even attempted to assess the outcomes and impact CRM had on the respective organizations and just five of the studies were related to military aviation. The low percentage of studies addressing results of CRM may be due to difficulty in linking the behavioral training to direct and tangible outcomes. Isolating the variables that contribute to the achievement of the mission is nearly impossible. Furthermore, even if a mission is considered a success, it does not necessarily imply efficiency or effectiveness.

An accepted method of analyzing CRM's organizational impact on aircrew coordination is to analyze historical safety data. This is typically accomplished through examining aviation mishap occurrences and rates over a period of time, usually years. Alkov (1989), in a study looking at naval aviation, cited a decrease in mishap rates as a result of CRM training across three specific categories of naval airframes (rotary, attack bombers, and multi-crew fighters). This study was performed over a four year period from 1986-1990. The estimated cost of the CRM program (over a five year period) was less than a million dollars, which is minimal compared to the loss associated with personnel or aircraft. A handful of other studies have evaluated safety mishap rates in naval aviation across the 1977-1996 timeframe. Chapter VI of this thesis presents a detailed look into these earlier studies with regards to mishap rates and uses them as a baseline to analyze CRM's impact on safety over the last decade.

D. SUMMARY

When evaluating any program, it is imperative to measure the training based on the organizational objectives. Evaluation is based on information gathered from measures at each criterion level including reactions, learning, behavior, and results. A successful evaluation does not solve all training issues, but it can shed light on the effectiveness of a program's content, delivery, and impact on the organization. There have been several studies on the success and implementation of Crew Resource Management training in military aviation. However, most of the studies were conducted some time ago. The research carried out in this thesis built upon prior studies and applies the principles of three levels of Kirkpatrick's topology to provide an up to date assessment of the naval CRM program. Prior CRM studies relevant to this analysis were broken down into the criteria of reactions, learning (in terms of both attitudes and knowledge), and results. Each level is discussed in more detail in the subsequent chapters.

THIS PAGE INTENTIONALLY LEFT BLANK

III. NAVAL AVIATION CRM REACTION ANALYSIS

A. BACKGROUND

Learner reaction is one common metric collected when evaluating a training program. It is the first level of assessment in Kirkpatrick's topology of training evaluation (Kirkpatrick, 1976). Reactions of individuals from training can be collected using a variety of techniques including questionnaires, interviews, focus groups, and observation (Phillips, 1996). Although many different methods exist to gather reaction data, feedback is frequently collected through the design and implementation of survey assessments. These reaction surveys are traditionally comprised of a paper-based questionnaire administered at the conclusion of a training curriculum. The surveys are normally structured for respondents to answer questions regarding their satisfaction using a Likert rating scale. One of the benefits to using paper-based surveys is that individuals' responses are generally collected immediately following the conclusion of the training. Recency effects on memory indicate participants are more likely to recall their feelings and experiences regarding a program right away (Wickens & Hollands, 2000, p. 281). Too much time elapsed between participation and eliciting feedback can increase assessment difficulty.

As the use of technology becomes more widespread in training, it has become common practice to distribute electronic or web-based satisfaction surveys. Electronic surveys also can be administered quickly after training, but risk losing oversight in timely completion of the assessment. A major benefit in distributing electronic surveys is having the ability to sample learners in distance learning programs. In addition, the data collection and analysis can be streamlined. To understand the value in assessing reactions of CRM training, one must consider the differences between enjoyment and perceived usefulness of the program.

1. Reaction Affect and Utility

Reaction feedback can include two distinct types of data indicating the satisfaction of a training program. These include (a) the participants' feelings, thoughts, or emotions about the design, content, and delivery of concepts in the program and (b) the perception of whether the training was worthwhile or useful. Alliger et al. (1997) referred to these two measures as affect and utility. Affective reactions are usually easier to assess as compared to utility. Evaluating trainees' affective reactions is as simple as having participants rate how they "liked" the training. This can be either an overall rating of the entire program, or specific to each topic or session.

The second type of measurement is a utility-based reaction. These questions are centered on whether a training program will translate well to the actual job or task. Utility assessments are designed to discover the perceived usefulness or value of the training from its participants (Alliger, Tannenbaum, Bennet, & Holly Traver, 1997). Helmreich and Wilhelm (1991) claimed that "it is axiomatic that training will have more impact if participants perceive it as useful." If an individual perceives the training is going to be utilized and provide value to their future experience, this can increase their willingness to participate and increase motivation.

Furthermore, it is entirely possible that reaction assessment data might include a combination of affective and utility measurements. But, do reactions have any bearing on the effectiveness of the training? Many would agree positive reactions do not necessarily link to an increase in knowledge or attitude changes. Affective reactions seem to have little, if any, relationship to amount of learning transfer. However, utility reactions are believed to influence motivation and willingness to learn (Salas, Burke et al., 2001).

2. Previous Studies of CRM Reactions

O'Connor et al. (2008) conducted a meta-analysis of 16 prior CRM studies which looked at both the affective and/or utility reactions of CRM participants. The studies represented CRM training programs in various industries including commercial and military aviation, medicine, nuclear power, offshore oil production, and commercial shipping. Of the 16 analyses, seven included an assessment of the reactions to the training. These seven studies reported participant responses using a 5-point Likert scale. The combined mean response across the seven studies indicated a general favorable reaction to CRM training (M = 4.18; O'Connor et al., 2008). Salas, Burke et al. (2001)

reviewed 58 studies evaluating effectiveness of CRM programs in a similar fashion. Of the 58 studies, 27 (46%) looked at participants' reactions, many of which indicated participants had a favorable reaction to the CRM training programs.

Other studies have addressed reactions related to specific components of CRM training. Results from some of these studies suggested CRM training had greater perceived utility when administered through scenario-based learning and case studies rather than traditional lecture formats (Schiewe, 1995). Positive affective reactions are believed to be amplified with training that promotes interaction rather than passive teaching. However, it should be noted that different industries are likely to have varying CRM training concepts and delivery methods. For example, a commercial aviation CRM strategy might focus more on crews who do not work together on a continuous basis, whereas CRM in the medical field may approach CRM team cohesion concepts with surgical teams who routinely work together. Nevertheless, each of these earlier meta-analyses indicates trainees, in general, sense the utility of CRM training. The literature concerned with reaction analysis exclusively for naval aviation CRM is discussed below.

3. Reactions to Naval CRM Training

Several studies have examined the reactions of naval aviators to CRM training. Five studies, in particular, looked at reactions and the usefulness of CRM training across various platforms of naval aviation. Each of the five studies indicated that aviators in both the Navy and Marine Corps tended to be enthusiastic about the usefulness of CRM training (Baker, Bauman, & Zalesny, 1991; Salas et al., 1999; Stout, Salas, & Fowlkes, 1997; Stout, Salas, & Kraiger, 1996). Many CRM participants felt the training was beneficial in aspects of flight safety, crew coordination, and mission accomplishment (Salas et al., 1999). The reactions from these studies were generally positive and consistent across different airframes. The studies included responses from rotary wing and fixed wing aviators.

4. Reaction Hypotheses

Despite numerous studies in aviation illustrating positive reactions towards CRM topics, the most recent collection of Naval aviation CRM reaction data was obtained well over a decade ago. As CRM training program for Navy and Marine Corps aviation has not drastically changed since that time, one might assume this would lead to observing the same positive reactions towards CRM. This section attempts to examine this assumption. Although the Navy's CRM program has remained relatively constant, it is possible reactions of CRM participants could change over time. Therefore, it is important to re-evaluate the reactions of naval aviators using more recent data.

This chapter assesses Kirkpatrick's first level of evaluation (reactions) for naval aviators. An examination of CRM reactions was accomplished by analyzing a sample of reaction data gathered from Navy and Marine Corps aviators who participated in CRM training over a period of nine years using data collected via surveys from fiscal years 2000 to 2008. This section of the CRM analysis used the data to evaluate three experimental reaction hypotheses. The first hypothesis analyzes general changes and trends of aviator reactions from Naval CRM training over the past nine years.

> ***H_1: Reactions of aviators to the Navy CRM training program have become less positive over time.***

The amount of exposure for all aviation crewmembers to CRM concepts and training could affect reactions over a period of time. As CRM training continues to be required on a regular basis, there is potential for the training to become repetitive and stale. If this were the case, it could reflect in the reactions of the trainees across several years.

The second proposed hypothesis evaluates differences in reactions amongst more senior and experienced naval aviators as compared to less-experience junior officers.

> ***H_2: Junior aviation officers will be significantly more positive in their reactions towards the concepts addressed in CRM training than senior aviation officers.***

One method to distinguish experienced aviators in the military is to break groups into rank divisions. Here, the separation was made between junior and senior aviators. In this study, "junior" aviators were defined as officers in the O1 to O3 ranks and the senior aviation officers were identified in the O4 to O6 category. As a rule of thumb, O1 to O3 officers have less than six or seven years of experience, where O4 to O6 ranks typically have over seven years experience. Again, the amount of exposure to CRM training concepts can possibly play a role in enthusiasm for the training. The junior aviators may have more to gain from the training than more experienced officers may. Conversely, senior officers may recognize the need to train towards non-technical concepts or desire more CRM instruction to refresh concepts if they have been out of the cockpit for some time.

Since there are numerous types of airframes used throughout the Maritime services, each aircraft can be sorted into three different categories. The first is the rotary wing category, namely helicopters. The second is tactical jet fighters, also referred to as TACAIR. The third category includes most other types of airframes and is labeled large or "big wing" aircraft. The big wing aircraft group encompasses most large, fixed wing aircraft.

The third hypothesis seeks to investigate differences amongst aviators trained in various types of Navy and Marine Corps airframes.

> ***H_3: Large wing and rotary wing aviators will be significantly more positive in their reactions to the concepts addressed in CRM training than tactical jet aviators.***

Predominantly, tactical jets are comprised of single seat aircraft. The Navy and Marine Corps have various single seat airframes (e.g., F/A-18s). Since crew resource management generally focuses on aspects of crew coordination and team-related competencies, single seat aviators might demonstrate a lack of enthusiasm for the program. Research has suggested that, historically, a larger portion of mishaps due to some form of CRM failure was seen in rotary wing aircrew compare to what was seen in TACAIR (Wiegmann & Shappell, 1999). This could be a result of having multi-person crews in rotary wing aircraft as compared to single seat aircraft (this will be discussed further in Chapter VI). Nevertheless, the perceived utility of CRM in multi-crew rather than single seat aircraft may affect reaction outcomes as a result of CRM training.

B. METHODOLOGY

1. Data Collection

Currently, U.S. Naval Aviation commands are directed to administer a command safety climate assessment approximately every two years (Naval Safety Center, 2009). These climate surveys are designed to gauge a unit's general safety posture and used primarily as a feedback tool for Commanding Officers. For aviation communities, each safety climate survey seeks to estimate the utility of the command's CRM program. In particular, one question from the safety assessment directly addresses with the CRM program and its perceived usefulness by trainees. The question attempts to investigate if a command's Crew Resource Management (CRM) program is helping to improve mission performance and safety. The question reads as follows: "My command's Crew Resource Management (CRM) program is helping to improve mission performance and safety." Each participant was asked to respond to the question by rating their level of agreement using a 5-point Likert scale (1 having answered "disagree strongly" to 5 having answered "agree strongly"). The responses to this item were examined as a measure of the fleet's reaction to the usefulness of CRM training.

2. Sample

CRM reaction responses were obtained from the Naval Safety Center from each naval aviation commands' safety climate surveys. For this study, the data collected included 51,570 responses from Navy and Marine Corps aviation officers who took a safety climate survey in the past nine years (2000-2008). The officers varied in seniority, ranging from the officer ranks of O1 to O6. Each aviator in the data set was classified into one of three categories based upon their airframe experience, self indicated at the time of the survey. The three categories included rotary wing (helicopters), tactical jets (TACAIR), and other large fixed wing (big wing) aircraft. A breakdown of the demographics for the officers is listed in Table 3. Each of the safety climate survey respondents remained anonymous and unidentifiable. Therefore, it was possible, even probable, aviators responded to different safety climate assessments on multiple occasions over the nine-year period. The assumption of sample independence should still hold true as responses were categorized by year and analyzed separately.

Table 3. Percent Representation of Reaction Response by Category (2000-2008)

	O1-O3	O4-O6	*Total % of sample*
Helo	26.49%	9.78%	36.27%
TACAIR	25.52%	11.89%	37.41%
Big Wing	15.67%	10.65%	26.32%
Total % of sample	67.68%	32.32%	

C. DATA ANALYSIS AND RESULTS

1. Overall CRM Reaction Trend

Reactions of U.S. Naval aviators to CRM training were analyzed by fiscal year to determine any change in reaction over time. Across all ranks and airframes the average response rating provided by the sample of Navy and Marine Corps aviators regarding the utility of CRM concepts from 2000 to 2008 was found to be generally positive (M = 4.09, SD = 0.70). The reaction responses indicate a very slight increase in positive reactions over the nine-year period (see Figure 2). The data was sorted by year and independently analyzed. Each year was compared the others in the group. A Kruskal-Wallis non-parametric analysis of variance (ANOVA) was used to examine differences in mean reaction responses to CRM training across the nine year period (χ^2 = 338.378, df = 8, $p<0.01$). It was found that there was a significant difference between the earlier years (2000-2002) as compared to mean reaction responses seen in later years (2003-2008). This can also been seen in Figure 2.

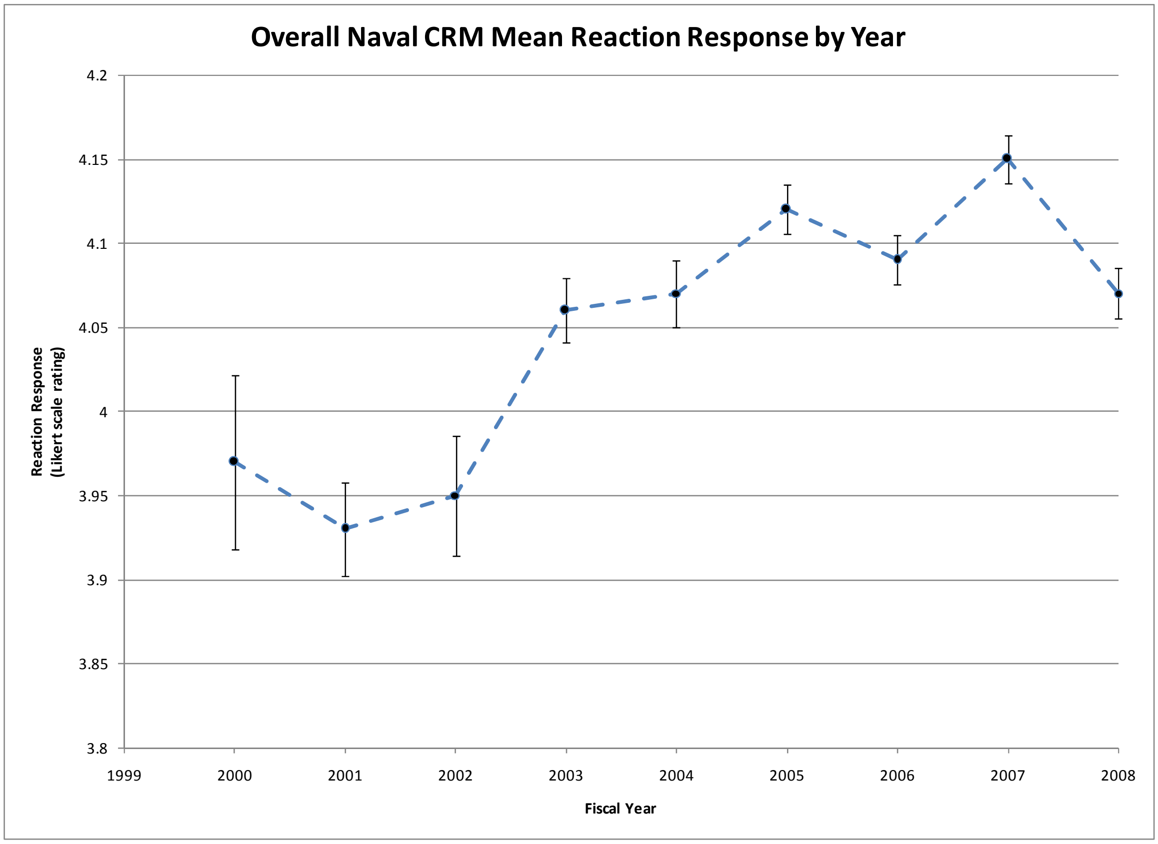

Figure 2. Mean comparisons of Naval aviation CRM reaction response from fiscal years 2000 to 2008 (with standard error bars included for each year).

2. Reactions Response by Experience

Navy and Marine Corps aviation officers were divided into two experience levels based on military rank (junior officers, O1 to O3; senior officers, O4 to O6). The average reaction responses to CRM training were examined by year and by rank. Figure 3 shows the graphical comparison of mean reaction responses for naval aviation officers by military rank across fiscal years 2000 to 2008.

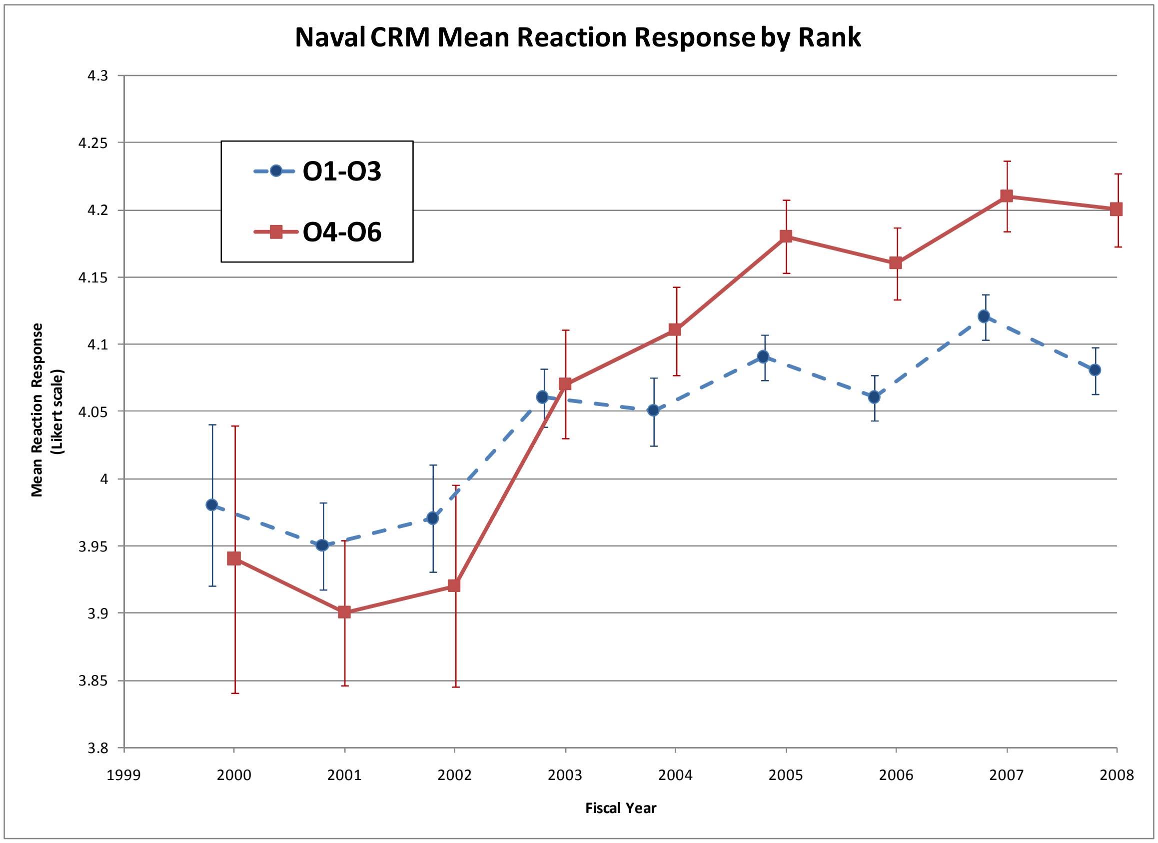

Figure 3. Mean comparison of Naval aviation CRM reaction response from fiscal years 2000 to 2008 by rank (with standard error bars included for each year).

Differences between the two rank groups were analyzed independently by year. A nonparametric statistical comparison using the Mann Whitney test was conducted to look at differences in rank independently for each year. The results of the Mann Whitney test are shown in Table 4. As can be seen, no significant difference was found from fiscal year 2000 to 2003. However, from 2004 to 2008 senior officers were significantly more positive in their reactions to the CRM training.

Table 4. Mann Whitney Comparison for Naval Reactions by Rank

Year	Sample (n)	*U*	p-value
2000	869	78363	n/s
2001	3184	1025008	n/s
2002	1737	278893	n/s
2003	4929	2181732	n/s
2004	4770	2472111	$p < 0.01$
2005	8423	6806289	$p < 0.01$
2006	8883	7507284	$p < 0.01$
2007	10176	9885232	$p < 0.01$
2008	8599	6835045	$p < 0.01$

3. Reaction Responses by Airframe

The data set was partitioned once again into reaction response based on each Navy and Marine Corps aviators' indicated experience with respect to airframe (Helo, TACAIR, Big wing). The yearly trend in mean CRM reaction response corresponding to aviator airframe is presented in Figure 4.

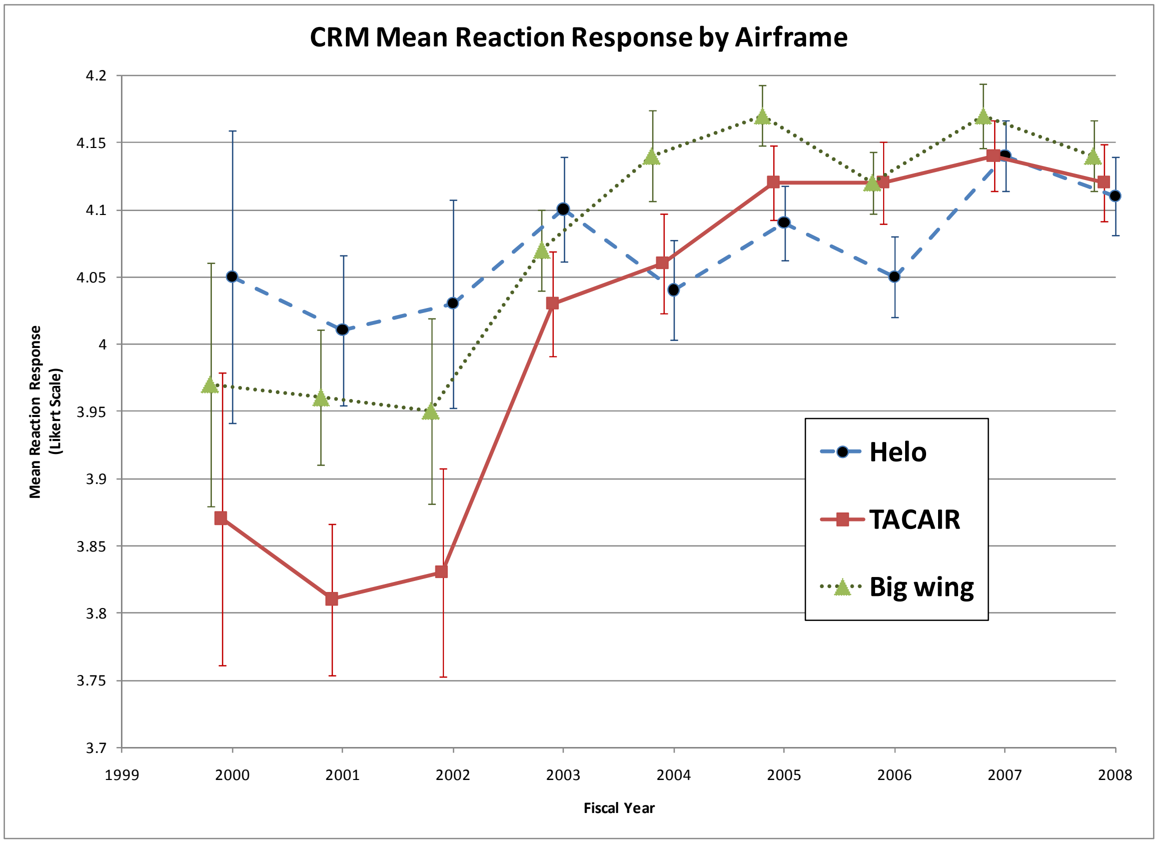

Figure 4. Mean comparison of Naval aviation CRM reaction response from fiscal years 2000 to 2008 by airframe (with standard error bars included for each year).

A supplementary statistical analysis was conducted to compare the means across each of the three airframe categories across the nine year time frame. Because the data was assumed to be a non-normal distribution, independent and non-parametric in nature, a Kruskal-Wallis one-way analysis of variance test was performed. The results of the Kruskal-Wallis test (df = 2) is shown in Table 5. As can be seen in Figure 4 and from the Kruskal-Wallis test, TACAIR aviators had significantly less positive reactions to CRM training in fiscal years 2001 and 2002. However, by 2003 TACAIR reactions became increasingly more positive and similar to those of rotary and big wing aircrew. The differences seen between the groups in 2004 through 2006 are due to a slightly higher mean reaction response of big wing aviators as measured up to the mean reaction response of TACAIR and rotary wing aviators.

Table 5. Kruskal-Wallis Comparison Using CRM Reaction Means by Airframe

Year	Sample (n)	$\chi 2$	p-value
2000	869	8.26	0.016
2001	3184	38.34	$p < 0.01$
2002	1737	19.76	$p < 0.01$
2003	4929	6.75	0.034
2004	4770	15.23	$p < 0.01$
2005	8423	17.83	$p < 0.01$
2006	8883	21.41	$p < 0.01$
2007	10176	7.01	0.03
2008	8599	6.35	0.042

D. DISCUSSION

1. General CRM Reactions

Hypothesis one (reactions of aviators to the Navy CRM training program have become less positive over time) was not supported. In fact, the reaction response increased very slightly over the period from 2000 to 2008. However, the change in mean response was not significant year by year. The overall results suggested that participants were positive about Naval CRM training and perceived the CRM training as useful.

The overall mean reaction response to the Naval CRM program (M = 4.09) indicates that CRM is generally well accepted. The results indicated a similar positive mean reaction response as compared to the previous CRM reaction studies in naval aviation that also showed positive reactions (O'Connor et al., 2008). These findings were also consistent with previous literature reviews concluding that CRM participants perceived the training as useful with favorable reactions (O’Connor et al., 2002; Salas, Burke et al., 2001; Salas et al., 2006). While the reactions, in general, were positive across all naval aviators, there was a considerable amount of variability in participant reactions. Some factors affecting the variability might have been due to differences in instructional techniques or alterations of CRM training at a command level. Instructors can also play a critical role in trainees’ enthusiasm for the course.

2. CRM Reactions by Rank

A slight increase in reaction response was seen in both junior and senior Navy and Marine Corps aviators across the nine-year time period. Interestingly, the reaction responses for both rank groups appeared to increase in the positive direction, specifically between fiscal years 2002 and 2003. No evidence was found to suggest a reason for the sudden boost in reaction responses after 2003. A speculative reason might be due to a change in culture as CRM becomes more accepted in the naval aviation community.

Examining the two rank subsets, senior aviators appeared to be consistently less positive over the 2000-2002 time frames. However, after 2003, the senior aviators' positive reaction responses rose to eventually match and supersede those of junior aviators. This result was contrary to what was expected in hypothesis two (junior aviation officers will be significantly more positive in their reactions towards the concepts addressed in CRM training than senior aviation officers). A possible explanation for these trends might include an increase in the perceived utility of CRM concepts for senior officers who have been out of the cockpit for some time. Senior aviators also might believe CRM training is a good method to refresh critical non-technical skills used in aircrew coordination. Additionally, as aviators in the Navy and Marine Corps become more experienced, team cohesion and communication concepts may be more openly received. The speculation is that maturity may foster a more favorable reaction to risk mitigation and safety issues as compared to those seen in less experienced aviators.

3. CRM Reactions by Airframe

There was an increase in mean response ratings across all airframes (Helo, TACAIR, and Big wing) over the past decade. A comparable spike in reactions was noticed between fiscal year's 2002 and 2003. However, the increase in response ratings appeared to be larger for the TACAIR aviators during this time. TACAIR aviators revealed a lower mean reaction response from 2000 to 2002 as compared to those of the rotary and fixed wing personnel. Although the findings did not denote a significant difference between reactions of TACAIR and rotary or big wing aviators, the slight trend of lower reactions in the early years provides some support for the third hypothesis (large wing and rotary wing aviators will be significantly more positive in their reactions to the concepts addressed in CRM training than tactical jet aviators). Moreover, after fiscal year 2003, the TACAIR aviators' reactions came closer to those of rotary and fixed wing personnel. Findings may reflect previous efforts to tailor command CRM training more toward specific TACAIR applications.

Consideration must be given to the idea that individual responses are accurate and correspond to a person's true feelings. Darby (2008) found that using Likert scales to ascertain feedback of students in training programs may not be as accurate as once thought. The results of Darby's (2008) study suggested responses to Likert scale questionnaires tended to skew more towards positive responses on virtually all measures of evaluation. It is possible naval aviators may be susceptible to a favorable response bias as CRM concepts are universally accepted as "positive" behavioral characteristics. An individual who responds that they did not perceive the training as useful may be translated into a feeling a disagreement with the principles of the CRM program. Another bias may result from the aviation culture itself. Naval officers may feel that negative responses indicate a need for change, and that change typically implies unwanted added responsibilities or duties.

E. SUMMARY

Even though a reaction analysis is usually the easiest form of assessment criteria, it is an important part of evaluating any training program. The results of this study suggest reactions to CRM training in Navy and Marine Corps aviation are favorable. Despite the differences in CRM training from command to command, aviators across various types of aircraft perceived it as useful and indicated it was helping to improve mission performance and safety. Historically, comparable studies similarly found reactions to CRM training in aviation have been positive (O'Connor et al., 2002; Salas, Burke et al., 2001; Salas et al., 2006). Naval CRM reactions are predominantly positive, and if there are any changes at all over the last decade, it would be in the slightly more positive direction.

Nevertheless, it is important to remember that positive reactions do not necessarily correspond to the overall effectiveness of a program (Kirkpatrick, 1976). Positive reactions to naval CRM training do hint at the validity and credibility of the program, providing a metric for higher-level management support. Because the safety climate surveys were distributed across multiple naval aviation commands, the affirmative reactions of the participants implied that core CRM concepts are well received. However, the reality is that reactions are only the first step in evaluating a training program and further levels of assessment need to be completed. The next two chapters consider the second level of Kirkpatrick's (1976) hierarchy framework deal with learning outcomes related to Naval CRM.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. NAVAL AVIATION CRM ATTITUDINAL ANALYSIS

A. BACKGROUND

The next two chapters discuss methods to assess learning outcomes of the Naval CRM program. This analysis applies Kirkpatrick's topology to look at the second level of evaluation: the learning outcomes from CRM training. As previously mentioned, learning can be divided into two main functions or goals: changing participant attitudes and improving the knowledge or skill of the trainee. This chapter's primary focus is on an attitudinal evaluation of the CRM program.

1. Attitudes of Aviators

Attitudes of trainees often are confused with reactions (discussed in Chapter III). However, it is important to understand the inherent differences between reactions and attitudes. As discussed, reactions essentially reflect the participants' satisfaction with the training and whether it is perceived as useful. Attitudes gauge the level of acceptance and willingness to apply the concepts and skills learned as a result of the training. Adoption of training concepts plays a significant role in learning and can have bearing on motivation and resulting behavior. Attitudes of trainees have been determined to be influenced by personality, culture, delivery of instruction, and dynamics of groups within the training (Helmreich & Wilhelm, 1991). The attitudes of Navy and Marine Corps aviators are empirically or conceptually related to resource management in the aircraft. Helmreich (1984) ascertained that only through modifying attitudes can we substantially change observable behavior. Research in the civil aviation industry also has shown that attitudes about the management of flight-deck resources are relevant to understanding error (Helmreich & Merritt, 1998) and to the quality of crew coordination (Helmreich, Foushee, Benson, & Russini, 1986).

2. CRM Attitude Assessment Tools

The most popular technique for assessing pilot attitudes towards the concepts taught in CRM training is through the use of questionnaires. Most of the previous evaluations of CRM have applied or adapted the already developed Cockpit Management Attitudes Questionnaire (CMAQ; Gregorich & Wilhelm, 1993) as well as the Flight Management Attitude Questionnaire (FMAQ; Helmreich & Merritt, 1998). The feedback of the CMAQ comes from self-reported indication of attitudes associated with CRM concepts. It incorporates three major scales including ‘communication and coordination’, ‘command responsibility’, and ‘recognition of stressor effects’. The one factor the CMAQ fails to consider is an attitude corresponding to the cognitive aspects such as situational awareness, decision making, and workload management (O'Connor et al., 2008). The CMAQ has 25 different statements describing attitudes drawing from earlier NASA research dealing with crew coordination (Helmreich & Wilhelm, 1991). Responses to the CMAQ and FMAQ are based on a five-point Likert scale as participants are asked to rate their level of agreement for each statement.

The CMAQ was developed initially to appraise aviator attitudes regarding “interpersonal components” of flight crew performance. The questionnaire also is used to associate these attitudes to crew behavior (Gregorich, Helmreich, & Wilhelm, 1990). The CMAQ is a well-established training evaluation and research tool developed to assess the attitudes of flight crew with regards to human factors issues. Not only has the CMAQ formed the basis of an attitude assessment for CRM programs in aviation, but has been utilized in a number of industries (e.g., nuclear power controllers, aviation maintenance, air traffic control, medicine, offshore oil production, and U.S. Navy divers; Flin et al., 2008). The FMAQ, a subsequent revision of the CMAQ, was developed to examine cross-cultural differences in attitudes of flight crews (Helmreich & Merritt, 1998). The FMAQ included 16 additional items to the CMAQ based on aviators’ values (Seva, Gutierrez, Duh, & Chong, 2007) and incorporated Hofstede’s dimensions of national culture (Hofstede, 1980).

Prior studies using the CMAQ questionnaire routinely found a positive shift in attitudes of trainees to the concepts taught in CRM programs (Helmreich & Wilhelm, 1991; Salas, et al., 1999). However, there have been studies that revealed certain types of people do not have positive attitude changes (Helmreich & Wilhelm, 1991). These groups are commonly referred to as “Boomerangs” or “Cowboys” and are labeled as such because these individuals rejected some of the basic CRM principles. Remedial training programs for these individuals were not shown to be a success (Irwin, 1991). Nevertheless, the CMAQ remains the most popular measure to assess attitudes and provides a useful baseline tool to evaluate attitudes of naval aviators. The advantage of basing a questionnaire on the CMAQ is its proven reasonable psychometric characteristics.

3. Attitudinal Hypotheses

The analysis section of this chapter drew on results of an attitudinal assessment based on an adapted version of the CMAQ to evaluate two experimental hypotheses. The first hypothesis addresses potential differences in attitudes of naval aviators trained on different aircraft platforms.

> ***H_1: Large wing and rotary wing aviators will be significantly more positive in their attitudes to the concepts addressed in CRM training than tactical jet aviators.***

Anecdotal evidence suggests that tactical jet pilots tend to take CRM training with less seriousness than aviators who fly in other platforms. As indicated earlier, this may be largely due to the fact many tactical jet pilots are single seat aircraft. Therefore, many might assume CRM training is better applied to aircraft with multi-crew features. As seen in previous interviews carried out amongst 36 U.S. Air Force F-16 pilots, more than 85% of the pilots had a neutral, mixed, or less than positive attitude to CRM (Karp, Condit, & Nullmeyer, 1999).

A second hypothesis investigates differences in attitudes between aviators based on rank. Identical to the reaction analysis in Chapter III, this study categorizes senior qualified officers as aviators who fall in the O4 to O6 ranking group and junior officers are qualified aviators in the O1 to O3 category who have previously participated in a naval CRM training program.

> ***H_2: Junior aviation officers will be significantly more positive in their attitudes towards the concepts addressed in CRM training than senior aviation officers.***

Issues of rank between personnel may play a greater role in the military flight crews, and this may be at odds with the concept of assertiveness taught in CRM training (Guzzo & Dickson, 1996). This finding was supported in a survey of 272 U.S Navy divers using an attitude questionnaire based upon the CMAQ. The study found that junior divers were significantly more sensitive to the effect of personal limitations on performance, and showed a significantly greater willingness to want to speak up than senior divers (O'Connor, 2007).

B. METHODOLOGY

1. Instrument and Development

For the present study, a version of the CMAQ was specifically developed for the U.S. Navy and Marine Corps to assess the general attitudes of Naval CRM participants. The questionnaire was customized explicitly for naval aviation using familiar terms and concepts presented in Naval CRM. A pilot questionnaire was distributed to a group of 20 experienced naval aviators for their review and feedback. The comments were collected and used to develop the Naval Aviator Human Factors (NAHF) questionnaire (see Appendix A). Similar to the FMAQ and CMAQ, the NAHF consisted of questions associated with aviators' attitudes towards specific concepts in CRM training. A total of 37 statements were considered for the development of the NAHF. However, 6 of the questions were removed as they did not pertain directly to attitudes towards CRM concepts, leaving 31 statements. The 31 statements were then categorized into five attitudinal composite subscales:

- **My stress**: consisting of 6 items (items 1, 3, 4, 7, 8, 14; see Appendix B).
  This scale emphasizes the consideration of- and possible compensation for- stressors in oneself (e.g., 'Even when fatigued I perform effectively during critical times in a flight').
- **Stress of others**: consisting of 6 items (items 2, 13, 16, 18, 19, 27).
  This scale emphasizes the consideration of- and possible compensation for- stressors in other team members (e.g., 'Aircrew should mention their stress or physical problems to other team members before or during a flight').

- **Communication**: consisting of 6 items (items 5,1 0,1 1, 17, 23, 32).

  This subscale encompasses communication of intent and plans, delegation of tasks and assignment of responsibilities, and the monitoring of crewmembers (e.g., ‘The specific roles and responsibilities of aircrew in an emergency are identified during the preflight brief’).

- **Command responsibility**: consisting of 9 items (items 6, 9, 28, 30, 31, 33, 4, 35,36).

  Command responsibility includes the notion of appropriate leadership and its implications for the delegation of tasks and responsibilities (e.g., ‘The Aircraft Commander should take physical control and fly the aircraft in emergency and non-standard situations’).

- **Rules and order**: consisting of 4 items (items 15, 22, 24, 29).

  This subscale is concerned with adherence to rules and procedures (e.g., ‘Written procedures are necessary for all in-flight situations”).

2. Participants

The NAHF questionnaire was distributed to a sample of Navy and Marine Corps aviators attending training at the Naval Aviation Schools Command in Pensacola, Florida. Participation in the study was strictly voluntary. Responses were obtained from a total of 364 U.S. Naval aviators. A total of 134 (37%) of the participants were junior officers (01-03) and 230 (63%) were senior level officers (04-06). Based on reported qualifications and experience, each aviator was classified into one of three airframe categories including: big wing, tactical jet, and rotary wing (see Table 6 for a breakdown of categories for respondents).

Table 6. Attitudinal Respondent Categorization.

Respondents	Big Wing	Tactical Jet	Rotary
Junior officers	43	41	50
	(11.8%)	(11.3%)	(13.7%)
Senior officers	60	89	81
	(16.5%)	(24.5%)	(22.2%)
Total	103	130	131

3. Measures

Data were gathered from participants' responses to the NAHF attitude questionnaire based on their ratings of agreement from the five-point Likert scale ranging from A (disagree strongly) to E (agree strongly). Each question based on the aforementioned five attitudinal categories was randomly distributed throughout the questionnaire. The corresponding responses to the statements were converted to a numerical value ranging from 1 to 5 (1 having answered "disagree strongly" to 5 having answered "agree strongly"). The participants were broken down by demographics into two categories for analysis: experience in the type of aircraft and their positional rank.

C. DATA ANALYSIS AND RESULTS

1. Psychometric Analysis.

An analysis of the psychometric properties of the NAHF was performed. Each of the 31 items was examined, looking at the reliability of the scales used. The skewness, kurtosis, and correlation between similar composite subscale statements were determined and analyzed. Firstly, the skewness and kurtosis of the individual items were examined. A total of nine response items displayed excessive levels of skewness or kurtosis (see Appendix B). Seven of the items were related to the subscale of 'communications', one for 'stress of others', and one for 'rules and order'. Excessive skewness and kurtosis indicate the data are not normal. Further investigation needed to be conducted to determine the consistency across the separate subscales.

An assessment was made of the internal reliability of the subscale factors using Cronbach's Alpha. Cronbach's Alpha is a tool that provides a coefficient of consistency with ranges in value from 0 to 1. A higher score corresponds to a more reliable scale. An Alpha of 0.7 or greater is typically regarded as the acceptable range for the reliability coefficient; however, lower thresholds are seen in the literature (e.g., Gregorich et al., 1990 reported alpha values of between 0.47 and 0.67 for the CMAQ). Indices resulting in lower scores do not necessarily make the findings invalid. In fact, low alphas may indicate the diversity of measured constructs. From Table 7, it can be seen that dropping items from both the 'my stress' and 'command responsibility' sub-scales led to an increase in the resulting Alpha levels. However, it was not possible to improve the reliability of the 'rules and order' sub-scale. Therefore, this sub-scale was dropped from further analysis.

Table 7. Cronbach's Alpha Values Corresponding to Each Subscale (before and after)

Sub-scale	α_{Before}	α_{After}
My stress (2 items dropped)	0.09	0.52
Stress of others	0.57	-
Communication	0.51	-
Command responsibility (3 items dropped)	0.03	0.54
Rules and order (Scale dropped; 4 items)	0.001	-

2. Between Groups Comparison

A two-way, between-subjects analysis of variance (ANOVA) was completed with airframe (big wing, tactical jet, and rotary wing) and rank (junior and senior officer) on the mean item score for each sub-scale. There were no significant main effects or an interaction for the 'my stress' subscale. Only the significant effects are reported. For 'stress of others' there was a significant main effect of airframe ($F(2,358) = 6.7$, $p < .05$). Rotary wing aviators were more sensitive to the stress of others than tactical jet aviators (see Figure 5). Further, there was a significant interaction between airframe and rank ($F(2,358) = 6.7$, $p<0.05$). For fixed wing and helicopter aviators, junior officers were more sensitive to the stress of others than senior officers. The opposite was true for tactical jet aviators. For the ‘communication’ subscale there was a significant main effect of rank ($F(1,358) = 8.7$, $p<0.05$). Senior officers had attitudes indicative of more open communication than junior officers (see Figure 6). For ‘command responsibility,’ there was a significant main effect of rank ($F(1,358) = 15.4$, $p<0.05$), as senior officers had more open attitude to involving the whole crew than junior officers (see Figure 7).

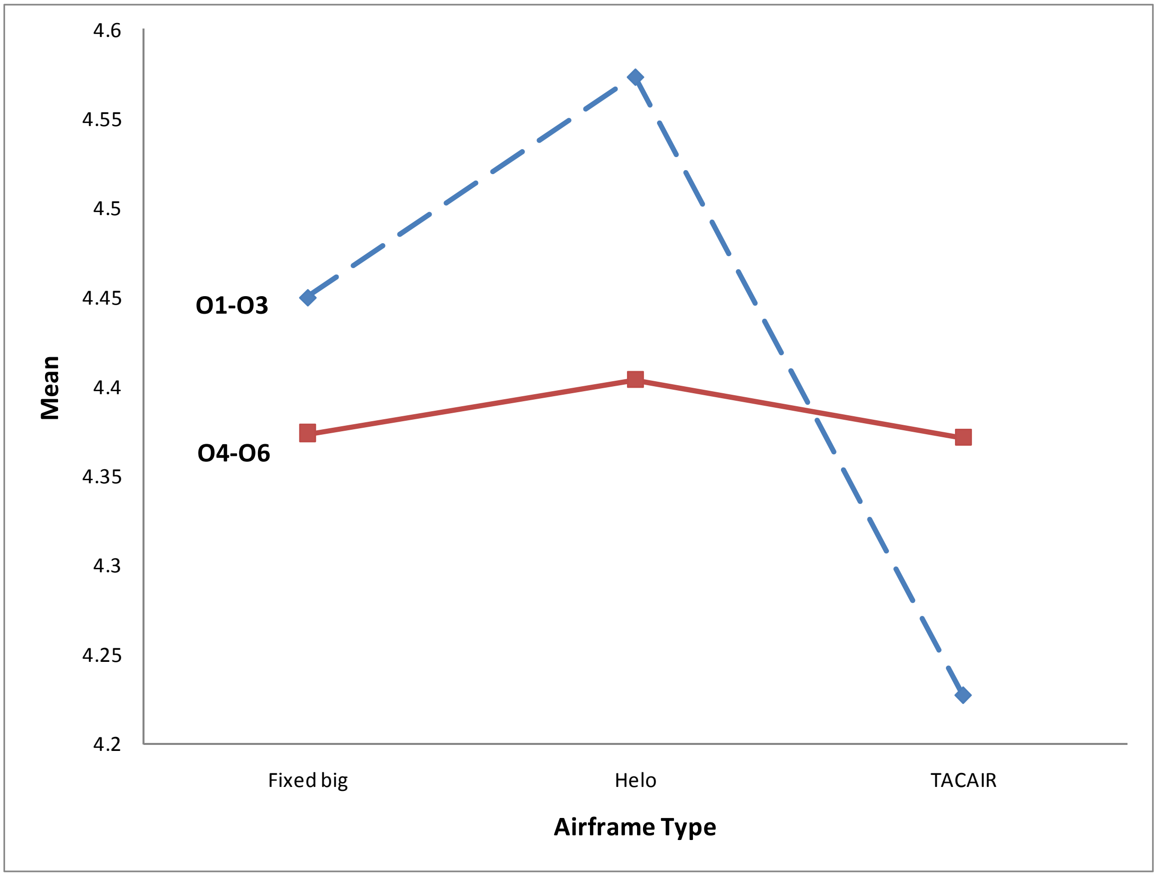

Figure 5. Mean factor score for ‘Stress of Others’ subscale for CRM attitudinal response as indicated on the NAHF questionnaire.

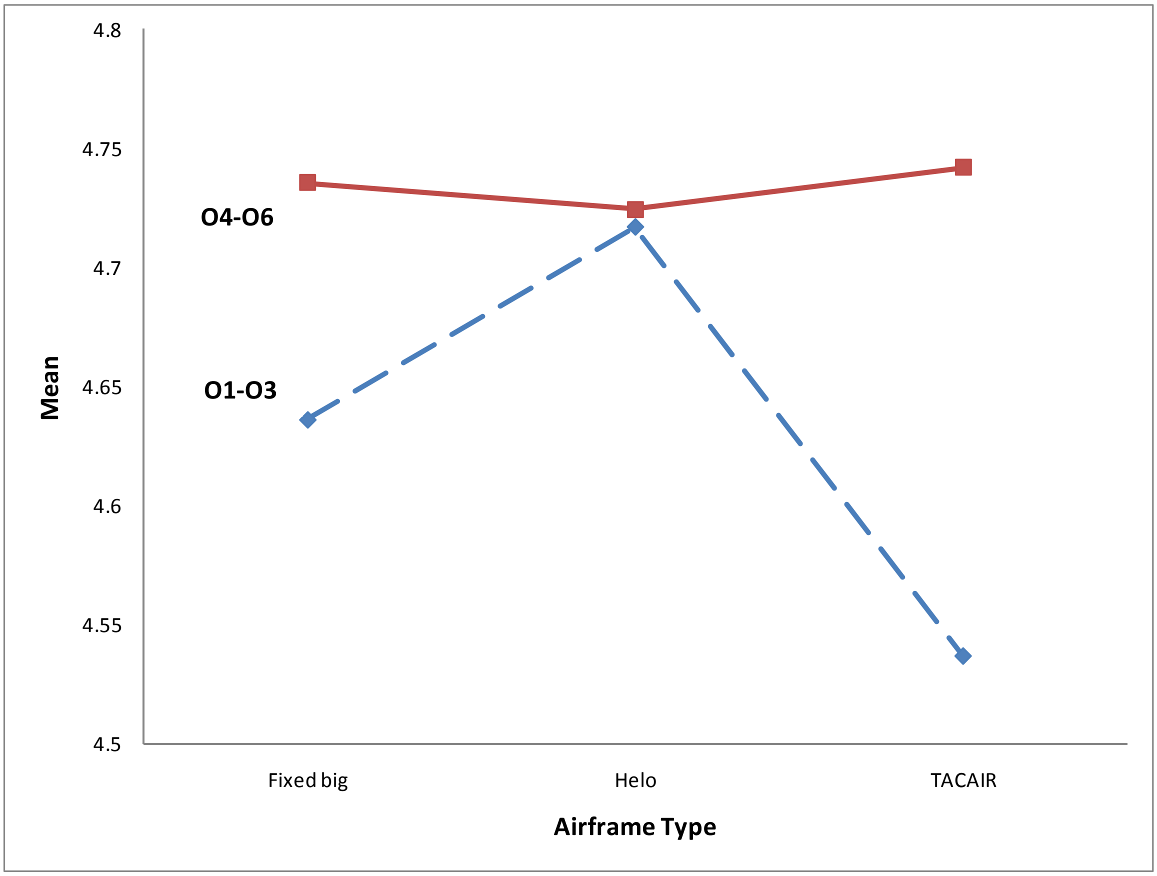

Figure 6. Mean factor score for ‘Communications’ subscale for CRM attitudinal response as indicated on the NAHF questionnaire.

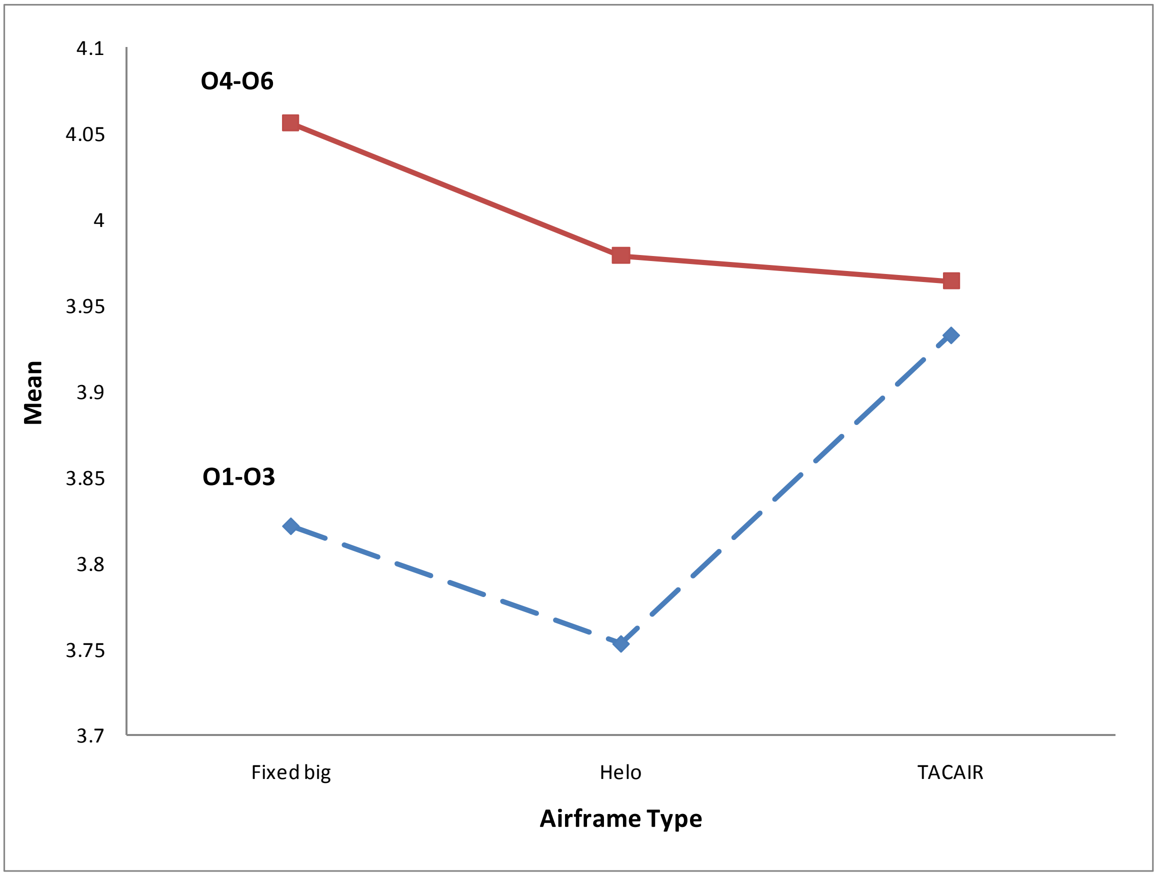

Figure 7. Mean factor score for ‘Command Responsibility’ subscale for CRM attitudinal response as indicated on the NAHF questionnaire.

D. DISCUSSION

1. Psychometric Properties

A total of nine items were discarded from the NAHF questionnaire due to excessive skewness or kurtosis. Discarding the statements also aided in an attempt to obtain acceptable Cronbach’s alphas. Discarding items is a common practice used in questionnaire development (DeVellis, 1991). The resulting Cronbach’s Alpha scores of the subscales were still relatively low after the nine items were dropped. The evaluation of the psychometric properties of the NAHF reported in this thesis is only preliminary. The recommendation would be to carry out a confirmatory factor analysis on the NAHF to assess the level of fit between the data and the proposed factor structure.

2. Between Groups Comparison

Hypothesis one (large wing and rotary wing aviators will be significantly more positive in their attitudes to the concepts addressed in CRM training than tactical jet aviators) was generally not supported. The only subscale for which there was a significant difference on the basis of airframe was ‘stress of others.’ Rotary wing aviators apparently were more sensitive to the stress of other team members than tactical jet aviators. These findings may reflect the larger crew and teamwork coordination required to fly a military helicopter (two pilots and an enlisted crew chief) versus smaller teams seen in tactical jets (either one pilot, or a pilot and navigator, or naval flight officer). However, even in a single seat tactical jet, the pilot is very unlikely to be flying without at least one wingman (a pilot in another aircraft performing the same mission). Therefore, the Navy’s CRM program, as applied in tactical aviation airframes, may benefit from an increased emphasis on the effect of the stress of others on performance.

The most interesting finding, and counter to the second hypothesis (Junior aviation officers will have significantly more positive attitudes towards the concepts addressed in CRM training than senior aviation officers), was that senior aviators seemed to be more supportive than junior aviators of an open cockpit climate as reflected by a significantly higher mean item score for the ‘command responsibility’ and ‘communication’ subscales. Research shows that high-performing teams have a climate of openness and trust, one where team leaders are receptive to alternative views and team members are unafraid to express their thoughts (Flin, 1996). The need for assertive behavior in junior team members has been sharply highlighted in aviation research (e.g., Foushee & Helmreich, 1988). The findings from the NAHF attitude questionnaire suggest that senior aviators are receptive to alternate views. However, the junior aviators may need to be persuaded that this is truly the case.

Although counter to what was expected, on reflection, this result may not be a complete surprise. The vast majority of the senior individuals who completed the questionnaire had been selected to be the Commanding Officer (CO), or Executive Officer (XO; second in command) of a squadron. Therefore, fostering a squadron with an open culture is within their interest as a leader. Further, many of these senior individuals had come back to aviation after spending a number of years 'out of the aircraft' performing other jobs. This may have led to a recognition that their flying skills were rusty and they would need to rely on the expertise of their junior personnel in the aircraft.

E. SUMMARY

The study described in this chapter provides a preliminary analysis of the attitudes of naval aviators towards several of the concepts addressed in CRM training. Recommendations for further attitudinal analysis should be considered. Additional research could be carried out to evaluate the fit of the data to the proposed factor structure. However, it is possible to draw two main conclusions from the attitude questionnaire analysis. Firstly, it was observed that tactical jet squadrons may benefit from training specifically targeted on recognizing, and addressing, the effects of stress in other aviators. Secondly, senior aviators appeared more supportive of an open cockpit climate than junior aviators. It is suggested that both of these issues should be addressed in recurrent CRM training.

Attitudes of naval aviators towards CRM concepts have repetitively shown to be well received and accepted. Yet according to Kirkpatrick (1976), an assessment of attitudes of trainees is only one aspect of the second level of a learning evaluation. To extend the evaluation of learning, knowledge transformations also must be examined. The next chapter discusses a method to analyze the other piece of a learning assessment by looking at the knowledge differences between different naval aviators.

V. NAVAL AVIATION CRM KNOWLEDGE ANALYSIS

A. BACKGROUND

1. Knowledge Evaluations

The second component in evaluating effects of learning, beyond evaluating attitudinal changes, is to assess the knowledge that was gained as a result of the training. According to Kirkpatrick (1996), in order to successfully gauge the amount of knowledge acquired, one must devise a method to measure the principles, facts and techniques learned from the program. Learning measures should be objective and quantifiable in nature and are not necessarily indicative of performance on the job (Goldstein & Ford, 2002). Performance on the job corresponds to behavioral changes, whereas knowledge acquisition addresses aspects of declarative and procedural information retained as a result of the training.

One accepted method used to assess learning is by administering examinations such as paper and pencil tests. Evaluating the effects of learning using paper and pencil tests is a fairly simple and quick way to estimate knowledge acquisition (Flin et al., 2008). Tests or exams can denote an increase of the knowledge of a program participant. Although exams are popular and a widely used method to assess learning, they also can be unintentionally misleading. For example, depending on the design of the test, the questions may be too hard (floor effect) or too easy (ceiling effect). Another way to evaluate learning of non-technical skills is through the use of vignettes or scenarios. The advantages of vignettes are that the measures can be unobtrusive (Kerlinger, 1996), easy to replicate, realistic, and interesting to the participant (Gliner, 1997). However, vignettes may produce unrealistic responses, become time consuming, and if not designed appropriately, difficult to understand (Flin et al., 2008). Analysis of vignettes can also lead to subjectivity and difficulty in quantifying the results.

2. Previous Knowledge Assessments of CRM

Despite the large number of studies evaluating CRM training, only a handful of studies reported a knowledge assessment. The most common techniques previously used to assess knowledge or skills from CRM training were multiple choice tests (O'Connor et al., 2002). A review of the literature indicated increases in knowledge of CRM participants as a result of the training. For example, O'Connor et al. (2008) conducted a meta-analysis of CRM evaluations, and found four studies that assessed knowledge gains of participants. Of these studies, three examined the effects of CRM training of U.S. Naval aviators with respect to knowledge acquired from the Navy's CRM course. Each of these studies used a multiple-choice test as the instrument designed to gauge the amount of knowledge change. Three of the four studies suggested participant knowledge of CRM concepts was greater after training, as compared to before (O'Connor, Hahn, & Salas, in press). In one study of 67 U.S. Naval aircrew personnel, Salas et al. (1999) found that knowledge of teamwork principles increased despite a lack of attitudinal changes. The aviators who received CRM training had shown higher scores (M = 12.6 out of 17) compared to a control group whose scores were slightly lower (M = 9.8 out of 17). Although many of these studies provided valuable metrics on CRM knowledge changes, all were conducted over a decade ago.

3. Knowledge Experimental Hypotheses

H_1: Senior naval aviation officers will score significantly higher on a CRM knowledge assessment as compared to junior naval aviators.

Since CRM training is directed to be conducted annually, senior aviators in the Navy and Marine Corps have had more exposure to training and experience with regards to CRM concepts. Therefore, general concepts addressing crew coordination and safety should be more familiar to senior naval aviators as compared to the junior aviators with less training and experience.

H_2: There will be a significant difference between aviators based on aircraft type (including rotary, TACAIR, and big wing).

The second hypothesis will determine the effect of airframe on knowledge scores. No differences are anticipated, despite CRM being administered uniquely from command to command. CRM teaches the same basic principles and non-technical behavioral skills throughout the organization and, therefore, should not have an effect due to airframe. If an effect should exist, further research would be warranted to determine inconsistencies in CRM instruction.

B. METHODOLOGY

1. Instrument and Development

A ten item multiple-choice exam was developed to assess knowledge acquired and retained with regards to concepts taught in the Naval Aviation CRM program. A retention test, known as the Knowledge Assessment for Naval Crew Resource Management (KANCRM), was developed for this thesis using a popular web-based survey assessment tool. Questions for the KANCRM were derived from material taught at the five-day Naval CRM instructor course. The questions of the KANCRM included a small sample of CRM topics and addressed specific knowledge and skills pertaining to human error, workload management, assertiveness, situational awareness, decision making, communication, mission analysis, fatigue, and stress effects on performance (see Appendix C). A pilot version of the assessment was distributed to non-aviation students at the Naval Postgraduate School for review and comment. Participants were asked to respond to each question and to indicate their best response by choosing from a list of provided answers.

The KANCRM was screened and approved by the Naval Postgraduate Institutional Review Board (IRB). All participants were provided an informed consent form, notifying them of their rights and risks as voluntary participants in responding to the questionnaire. No compensation was provided to those willing to participate.

2. Participants

The KANCRM was distributed to U.S. Navy and Marine Corps aviation officers currently attending either the Naval Postgraduate School in Monterey, California or the Naval Aviation Safety School in Pensacola, Florida. At the Naval Postgraduate School, the KANCRM was delivered a link to the web-based questionnaire via electronic mail to 101 U.S. Navy and Marine Corps pilots and Naval Flight Officers (NFOs). The response rate of the KANCRM for the Naval Postgraduate School was roughly 70%. In addition, the questionnaire was issued to a class of aviation officers attending training at the Naval Aviation Safety School in Pensacola. An additional 61 participants completed the assessment at the Navy Safety School; however, nine of these responses were redacted as they were not currently in the Navy or Marine Corps aviation field. The combination of the two samples resulted in responses from 123 naval aviators.

The participant demographics consisted of a wide range of experience and backgrounds. The U.S. Navy and Marine Corps aviators ranged in rank from O1 to O6 and had between 6 months and 21 years of flying experience. Navy aviators accounted for 67.5% of the total sample, as compared to 32.5% Marine Corps aviators. The participants consisted of 118 male and five female fleet aviators. Reported flight hours varied from as little as 120 hours to as much as 3,800 hours of flight time. Of the respondents, 65% indicated they were pilots and 35% NFOs. The sample represented 35% TACAIR, 35% rotary, and 27% big wing aviators (3% remained unidentified). Participant responses from the KANCRM were collected and analyzed based on the answers to the 10 questions and the results are discussed in the following section.

C. RESULTS

1. Differences in Rank versus Airframe

The overall mean correct score of the KANCRM for the sample was found to be 6.66 out of a possible 10 with a standard deviation of 1.36. Although a significant difference between the ranks was not observed, the junior naval aviators scored slightly higher (M = 6.71, SD = 1.25) as compared to the senior aviators (M = 6.62, SD = 1.68) as seen in Figure 8. Additionally, slight differences, although not found to be significant, were observed when mean scores were compared by airframe. Rotary wing aviators exhibited the highest mean score (M = 6.79, SD = 1.52) followed by TACAIR aviators (M = 6.65, SD = 1.36) and big wing aviators were the lowest (M = 6.49, SD = 1.12). However, there was a high degree of variability observed in all sets of aviation groups (see Figure 9).

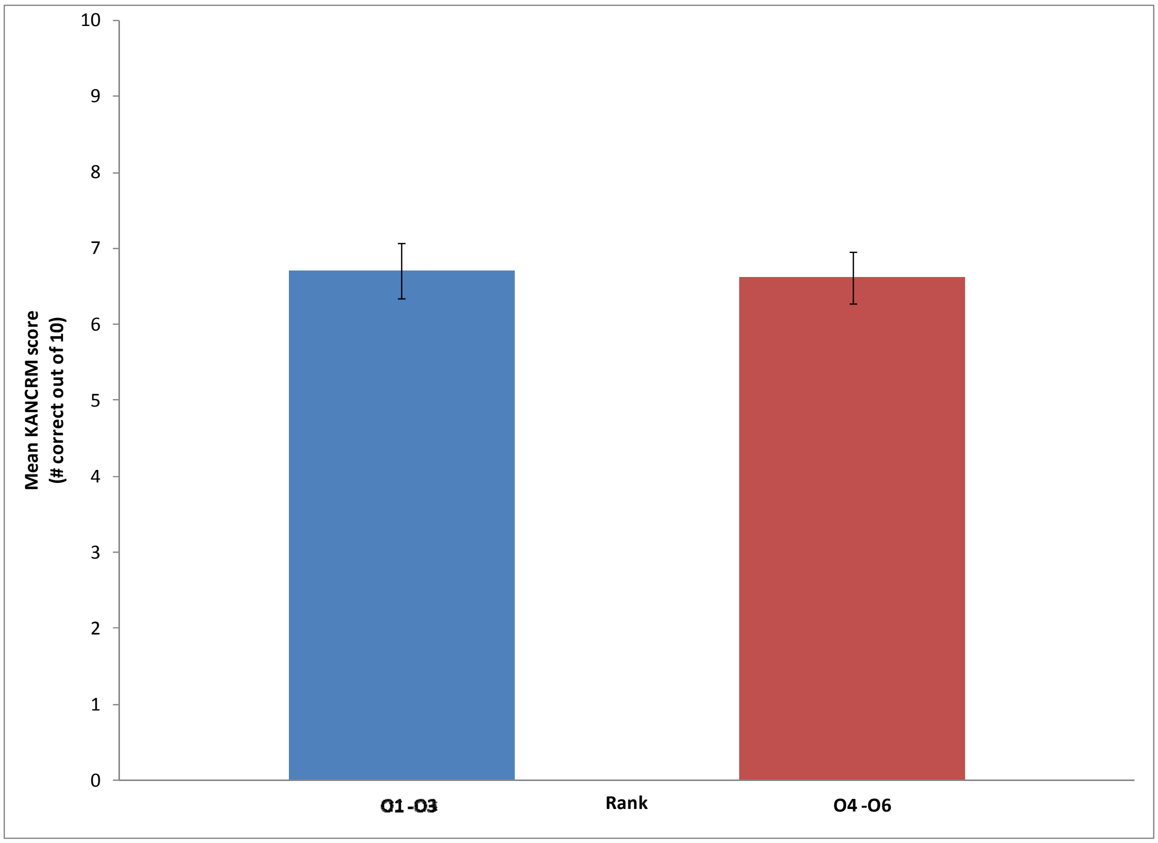

Figure 8. Mean KANCRM scores of U.S. Naval aviators by rank (with standard error bars). Scores are reported in number of correct responses out of 10 questions.

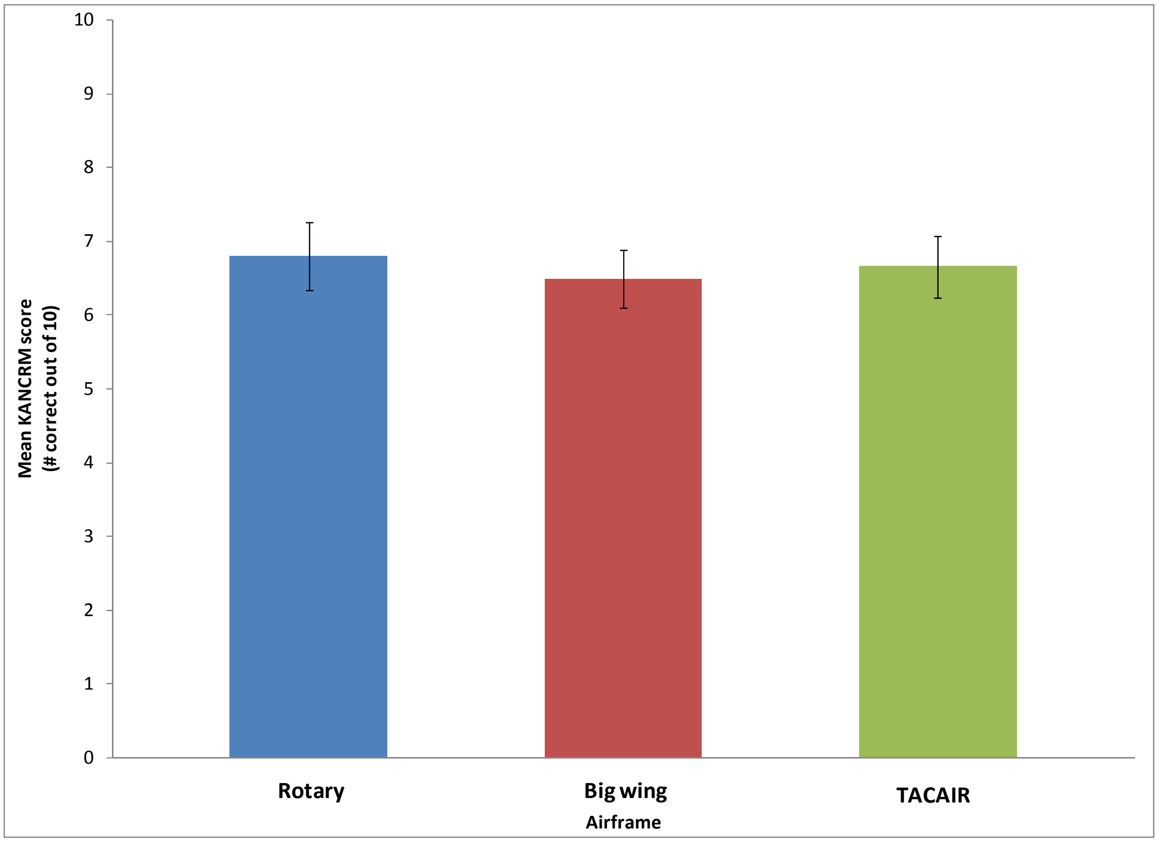

Figure 9. Mean KANCRM scores of U.S. Naval aviators by airframe (with standard error bars). Scores are reported in number of correct responses out of 10 questions.

2. Tests of Between Factors Effects

Further analysis was conducted to examine the effects between rank and type of airframe. A Levene's test of equality of error in the variances was calculated with no significant inconsistencies found (F(5,113) =1.90, p>0.05). In addition, a univariate analysis of variance (ANOVA) was conducted to test the effects of rank and type of airframe on CRM participant score. The ANOVA showed no significant difference between the groups or interactions (F(5,113) = 1.05, p>0.05). Even though no differences in mean scores of naval aviators were found to be statistically significant, senior rotary aviator scores were slightly lower than their junior rotary counterparts, while the senior TACAIR aviations scored higher than junior TACAIR aviators (see Figure 10).

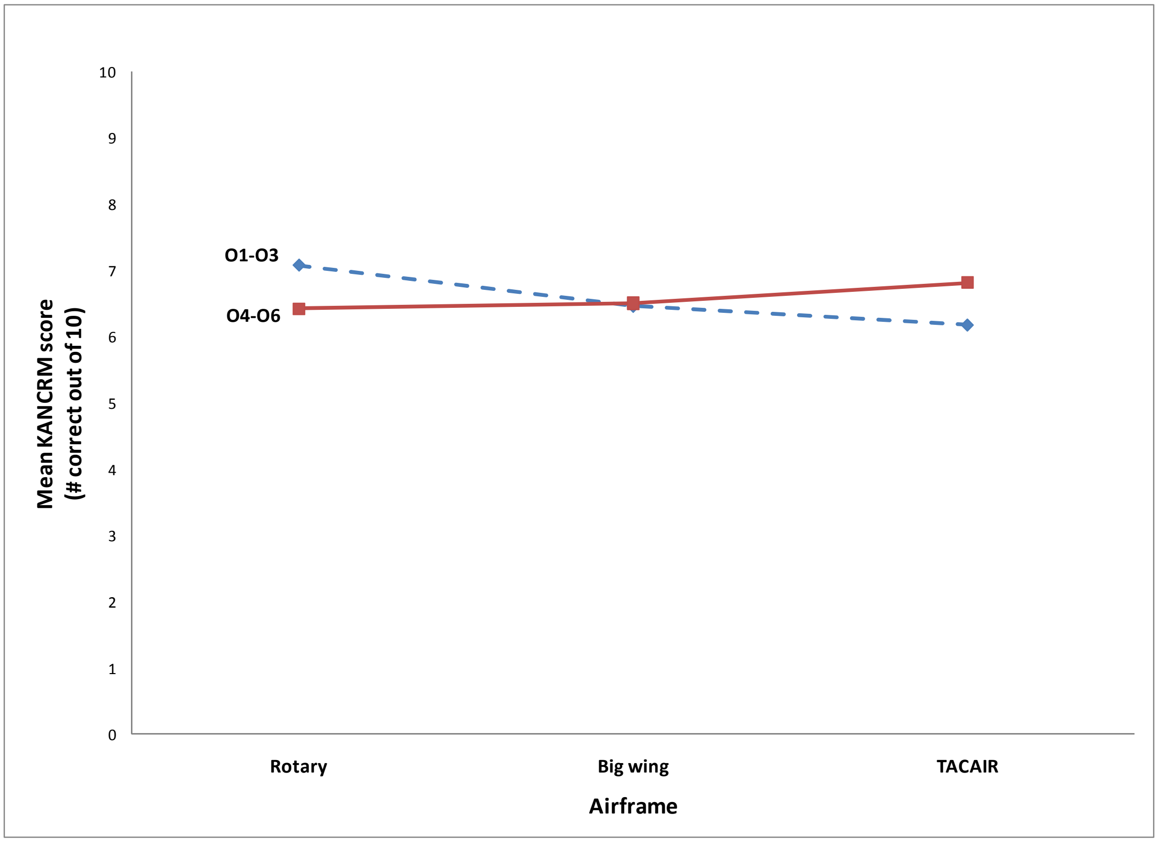

Figure 10. Comparison of mean KANCRM scores by rank across indicated airframe type. Junior officers (O1-O3) are represented by diamonds with a dashed line. Senior officer scores are represented with squares and a solid line (reported scores in number of correct responses out of 10 questions).

D. DATA ANALYSIS AND DISCUSSION

Hypothesis one (senior naval aviation officers will score significantly higher on a CRM knowledge assessment as compared to junior naval aviators) was not supported. Mean scores of both junior and senior aviators on the KANCRM were fairly consistent. Contrary to expectations, senior aviation officers failed to score significantly higher than their junior counterparts. It was assumed senior officers have encountered more experience and exposure to concepts addressed in CRM training. One possible explanation for the lack of difference could be a result of senior officers having been out of the cockpit for quite some time. Although difficult to ascertain from these findings, junior officers may have scored higher than expected if the questions, were in fact, tougher than anticipated. Non-technical skills and behaviors taught in CRM training also may be emphasized early in flight program instruction.

The second hypothesis (there will be a significant difference between aviators based on aircraft type) also was not supported, as expected. The results indicated no significant differences between mean scores of rotary, big wing, or TACAIR airframes. Lack of significance might indicate a consistency in CRM training. Similarity between the three types of aircraft could be a result of the large amount of variability seen in the scores across all airframes. The limited number of assessment questions also may have contributed to the high variability. However, some minor differences were noted and raise interesting questions. For example, rotary aviators, to some extent, scored a little higher as compared to TACAIR and big wing aviators. And, even though big wing aviators had the lowest mean score of the three, this difference could have been attributed to the sample (only having 27% big wing aircraft represented compared to 35% representation for rotary and TACAIR). The equivalent knowledge assessment scores point toward a positive outcome: CRM concepts might be stressed equally across all Navy and Marine Corps aviation communities.

Another interesting result was the overall mean score seen across all participants ($M = 6.66$ out of 10, $SD = 1.36$). This average low score could have a variety of explanations. The assessment questions may have been poorly designed and included awkward wording. This could have resulted in a potential floor effect for certain types of questions (either too difficult or ambiguous). Another explanation might be CRM training programs are not equally emphasizing certain non-technical concepts. Investigation into each question revealed the questions with the lowest mean scores addressed human error in mishaps, situational awareness, and errors in decision making. Since there were only 10 questions with limited topic coverage, a low score is a potential consequence. Further studies may want to consider adding additional questions in all CRM concept areas to evaluate any differences.

E. CONCLUSIONS

Knowledge assessments of CRM training programs are not routinely conducted. In Naval CRM training, tests are not distributed to evaluate participants' understanding and retention of the concepts addressed. Rather, completion of the CRM course fulfills the requirements set forth by the Department of the Navy (Chief of Naval Operations, 2001). Evaluation of training in terms of learning is an imperative part of reviewing whether a training program is meeting its intended objective. Gauging whether a training program is being culturally accepted by all elements of an organization also is important to assess. This section contributed to the validation of the Navy's CRM program in terms of knowledge of CRM concepts between different ranks and communities within naval aviation. The results supported the Naval CRM program's success in terms of similar knowledge gained from training across each aviation community. The reactions and learning from a training program are only truly observed in the resulting behavior of the trainee; however, a in depth analysis of behavior is beyond the scope of this research effort. Still, there is one additional level of Kirkpatrick's (1976) hierarchy of evaluation to consider: the training's impact on the organizational and its results.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. NAVAL AVIATION CRM RESULTS ANALYSIS

A. BACKGROUND

The fourth and final level of Kirkpatrick's hierarchy of training evaluation is analyzing the results or impact on an organization. This is the highest level of evaluation as it gauges how effective a program is in meeting the stated desired outcomes in the organization. The bottom line is often seen in the results. Many organizations rely heavily on this metric to evaluate their training programs. Results can be, in many cases, complicated by a variety of factors and extremely difficult to assess (Phillips, 1996). In most cases, the challenge arises from identifying the causal relationship between the training process and the end product. For example, in aviation it is nearly impossible to attribute any one failure as the single source of an accident. Most mishaps are caused by, or attributed to, two or more inherent factors. A dynamic interaction exists between mechanical, environmental, and human systems and each one has multiple factors that have an effect or impact on the others. All the same, consideration must be taken to address which factors influence the outcome. Identifying the factors of a mishap is accomplished through logical analysis. This section of this thesis is devoted to an examination of naval aviation mishap rates over the past decade to understand outcomes related to CRM training.

1. Naval Mishap Classifications

The Department of the Navy (DON) defines a naval aviation mishap "as an unplanned event or series of events, directly involving naval aircraft or UAVs" which result in $20,000 or greater cumulative damage to the aircraft or property, or where the outcome directly leads to injury of personnel (OPNAVINST 3750.6R, Department of the Navy, 2007). Each mishap is rendered a classification based on the severity of the mishap. Figure 11 illustrates the Navy's classification system of aviation mishaps, placing them into categories A, B, or C respectively.

	SEVERITY CLASS		
MISHAP CATEGORY	A	B	C
FLIGHT MISHAP (FM) Intent for flight existed, and $20,000 or more DOD aircraft/UAV damage occurred.	Total damage cost is $1,000,000 or more and/or aircraft destroyed and/or fatal injury and or permanent disability	Total damage cost is $200,000,000 but less than$1,000,000 and/or permanent partial disability and/or hospitalization of three or more . personnel.	Total damage cost is $20,000 but less than $200,000 and/or five lost workdays injury.
FLIGHT-RELATED MISHAP (FRM) Intent for flight existed with less than $20,000 DOD aircraft or UAV damage	Total damage cost is $1,000,000 or more and/or fatal injury and or permanent disability	Total damage cost is $200,000 but less than $1,000,000 and/or permanent partial disability and/or hospitalization of three or more personnel.	Total damage cost is $20,000 but less than $200,000 and/or five lost workdays injury.
AVIATION GROUND MISHAP (AGM) No intent for flight existed.	Total damage cost of $1,000,000 or more and/or aircraft destroyed and/or fatal injury and/or permanent total disability.	Total damage cost is $200,000 but less than $1,000,000 and/or permanent partial disability and/or hospitalization of three or more personnel.	Total damage cost is $20,000 but less than $200,000 and/or five lost workdays injury.

Figure 11. DON Naval Aviation mishap classification matrix (From OPNAVINST 3750.6).

2. Linking CRM to Mishaps and HFACS Identification

In U.S. Naval Aviation, the goal of CRM remains the same: to minimize crew preventable errors thus improving safety. Essentially, the better the use of all available resources, including people and equipment, the less chance a mishap will occur. CRM concepts focus on the non-technical behavioral skills required for effective team communication and coordination between aircrew personnel. Since studies have shown that the majority of accidents are related to human factors, CRM became the main approach to decrease the number of human related errors. In order to better understand the factors involved in an aviation mishap and identify human factors contributions, a post-mishap analysis must be performed.

Once a naval aviation mishap has occurred and is classified, further investigation is carried out to determine the casual factors involved. Answering questions as to what occurred that resulted in the accident is relatively straightforward. The challenging part of causal determination comes when assessing the factors and reasons of why a mishap took place and what conditions were present that led to the mishap. Following Reason's (1990) concept of latent and active failures, a classification system was developed to identify explanations of why mishaps occur (Shappell & Wiegmann, 2000). The product resulted in the formation of the Human Factors Analysis and Classification System, also known as HFACS. The HFACS taxonomy describes four levels of failure: 1) Unsafe Acts, 2) Preconditions for Unsafe Acts, 3) Unsafe Supervision, and 4) Organizational Influences. Similar to the "Domino Theory" or "Swiss Cheese Model", the HFACS system is based on the premise that casual factors arise from a combination, or series, of events or preconditions that lead to an accident. HFACS provides a method to further differentiate reasons behind why mishaps occur and is beneficial in linking CRM concepts and training to impact of flight safety.

3. Mishap Rates in Naval Aviation

Improvements have been made in commercial and military aviation to reduce the number of accidents and enhance safety. Naval aviation has been no exception. Numbers have shown that over the past fifty years, naval aviation has seen a noticeable reduction in the mishap rate from around 55 mishaps per 100,000 flight hours to less than four mishaps per 100,000 flight hours (Shappell & Wiegmann, 1996). This progress has been attributed to improved aircraft design, maintenance, materials, procedure standardization, and other safety programs. Although accident reduction is believed to come from a combination of decreased mechanical and environmental factors causes, human error has not decreased at nearly the same rate. Where mishaps due to mechanical and environmental causes have dropped to insignificant amounts in the past several decades, mishaps as a result of human error have only diminished 50% over the same time period (Wiegmann & Shappell, 1999).

4. Previous Studies Analyzing Naval Aviation Mishap Rates

Three noteworthy studies have been conducted looking at the trends in the mishap rate for naval aviation. The first study looked at U.S. Naval aviation mishap rates from fiscal year 1977 to 1992. The study distinguished between accidents involving single-piloted versus dual-piloted aircraft (Shappell & Wiegmann, 1996). Over the 16-year period, the rate of Class A, B, and C aviation mishaps decreased significantly, revealing a trend toward safer flight evolutions. However, the rate of decline was not found to be equal involving the percentage accidents attributed to mechanical and environmental issues as compared to those related to human factors error. In fact, the frequency of mishaps accredited to mechanical or environmental problems had dropped to insignificant amounts as early as the mid 1990s. Conversely, the frequency of naval aviation mishaps due to human error over the same time frame only fell by about 50%.

A similar trend was determined to exist in both single- and dual-piloted aircraft. The ratio of human related error to mechanical/environmental causes in 1977 was roughly 1 to 1, whereas in 1992 the ratio increased to approximately 9 to 1 for single-piloted aircraft and 12 to 1 for multi-crew aircraft (Shappell & Wiegmann, 1996). This is a significant increase in the proportion of mechanical/environmental to human attributed mishaps. Shappell and Wiegmann (1996) concluded the study by recommending a more focused effort to reduce the rate of mishaps in naval aviation despite the implementation of CRM related training programs.

A second major naval aviation mishap study reviewed mishap trends and causal factors over a four year period from 1986 to 1990 (Yacavone, 1993). This study included only Naval Class A mishaps in its final evaluation. Yacavone (1993) found that 59% of Navy and Marine Corps Class A aviation mishaps were attributed, in part, to aircrew human factors. Further analysis revealed that 45% of all naval aviation mishaps were accredited to "lack of aircrew coordination," the most common human factor causal error. These findings further sparked the Navy's interest in developing a standardized Aircrew Coordination Training (ACT) program that ultimately grew into the current Naval CRM program of today.

A follow on study was conducted in 1999 by Wiegmann and Shappell that investigated naval aviation mishaps from 1990 to 1996. This study examined both rotary and tactical jet (TACAIR) Class A naval mishaps over the six-year period to assess the role of human error and crew resource management failures. The results revealed that over 75% of the mishaps were attributed to human error and that 56% involved at least one failure related to CRM concepts (Wiegmann & Shappell, 1999). TACAIR mishap rates during this period were higher than those of rotary wing mishap rates. TACAIR mishap rates across the seven-year time frame were found to be approximately 4.38 mishaps per 100,000 flight hours as compared to the 2.98 mishaps per 100,000 flight hours for rotary aircrews. The authors noted that TACAIR operations are conventionally more hazardous and include more aggressive maneuvers than those of rotary aircraft. Despite the fact failures due to human error were fairly consistent across both airframes, rotary wing aircrew mishaps had a higher percentage associated with CRM failures than TACAIR (80.43% compared to 46.49%). Although the overall mishap rates remained relatively low during this time period, CRM failures continued to be a major causal factor involved. Since this was the most recent study of CRM's effect on naval aviation mishap rate, this thesis continued to build upon this research by evaluating Class A mishaps in naval aviation to examine the effects of CRM training in terms of the impact on the organization over the past decade.

5. Results Hypothesis

Wiegmann and Shappell's (1999) findings illustrated naval aviation's consistent mishap rate across fiscal year's 1990 to 1996. During those years, CRM became more widely accepted in both TACAIR and rotary wing communities. The impact of CRM training on the Navy as an organization was analyzed using mishap data obtained from the Naval Safety Center to investigate the following hypothesis:

> ***Crew Resource Management training has had a positive effect on the U.S. Navy and Marine Corps aviation Class A mishap rate in both fixed wing and rotary platforms.***

B. METHODOLGY

1. Objective

In this study, an examination was conducted of Naval Aviation Class A mishap reports (from 1997 until 2007) obtained from the Naval Safety Center. This evaluation considered the mishap rate in terms of both individual and aircrew error as indicated by the data. The mishap data was separated in terms of error associated with CRM and non-CRM failures that were identified as causal to the mishap. This assessment seeks to determine what fraction is likely resulting from CRM training and if CRM has had a positive effect on mishap rates over the past decade.

2. Safety Mishap Data 1997-2007

Naval aviation mishap data from fiscal year 1997 to 2007 was separated into three categories based on type of aircraft. The three categories included rotary, TACAIR, and a third category for the V-22 Osprey (a tiltrotor aircraft used in the Marine Corps and Air Force). The data did not include mishap data for large, fixed wing aircraft. Due to the V-22 being a newly introduced aircraft within the time frame being analyzed, as well as the high mishap rate during its evaluation period, these data were excluded from the analysis. Discarding the V-22 mishap data left two remaining airframes for comparison: rotary wing and TACAIR aircraft. The data included the number of mishaps attributed to human error as well as those attributed to at least one CRM-type failure. Each fiscal year included the corresponding number of flight hours for each airframe allowing for the determination of the yearly mishap rate. The results were analyzed and are discussed in the following section.

C. DATA ANALYSIS AND RESULTS

1. Overall Naval Aviation Mishap Frequencies and Rates

Over fiscal years 1997 through 2007, a total of 238 naval aviation mishaps were recorded (divided into 140 TACAIR and 98 rotary wing aircraft). The total documented flight time for both airframes was 8,363,663 hours. The distribution of flight hours was fairly consistent with 4,019,045 TACAIR flight hours and 4,344,618 flight hours in rotary aircraft. This resulted in a combined flight mishap rate across the time period of 2.85 mishaps per 100,000 hours of flight operations. Over the 11-year span, the TACAIR community had a slightly higher mishap rate of 3.48 mishaps per 100,000 flight hours as compared to 2.26 mishaps per 100,000 flight hours for rotary wing aircraft (see Figure 12). The peak mishap rate for TACAIR was reported in fiscal year 2003 with a rate of

5.15 and the lowest was 2.25 in 2007. For rotary wing aircraft, the maximum mishap rate was 4.32 in 2002 preceded by the lowest mishap rate of 1.01 in 2001.

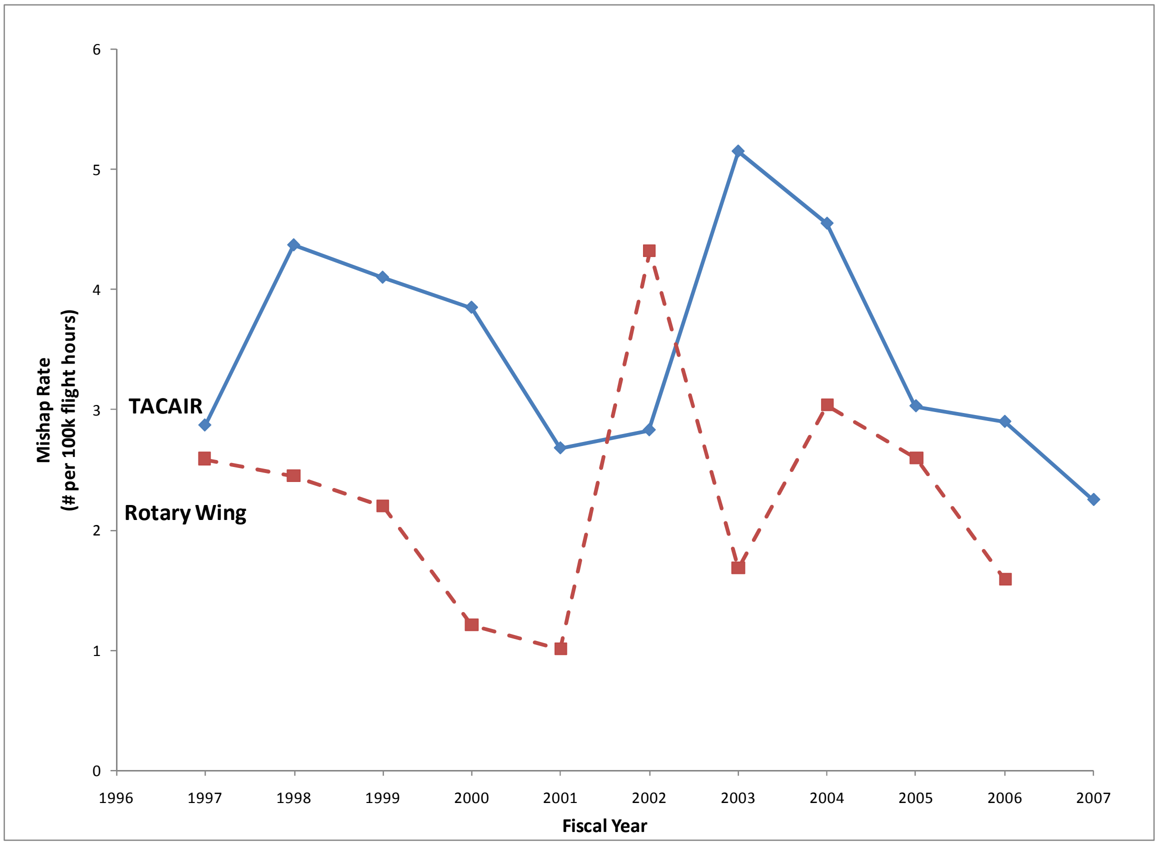

Figure 12. Mishap rates for TACAIR (solid line) and rotary wing (dashed line) aircraft across fiscal years 1997-2007.

2. Mishaps Related to Human-Error

Of the 238 mishaps in naval aviation from fiscal year 1997 to 2007, 90.3% were attributed to some form of human error. The relative percent of mishaps for both the TACAIR and rotary wing communities were comparatively equal. The TACAIR data shows that 90.7% of the 140 mishaps were attributed, at least in part, to human error; whereas rotary wing aircraft had 89.8% of the 98 mishaps ascribed to human error (see Figure 13).

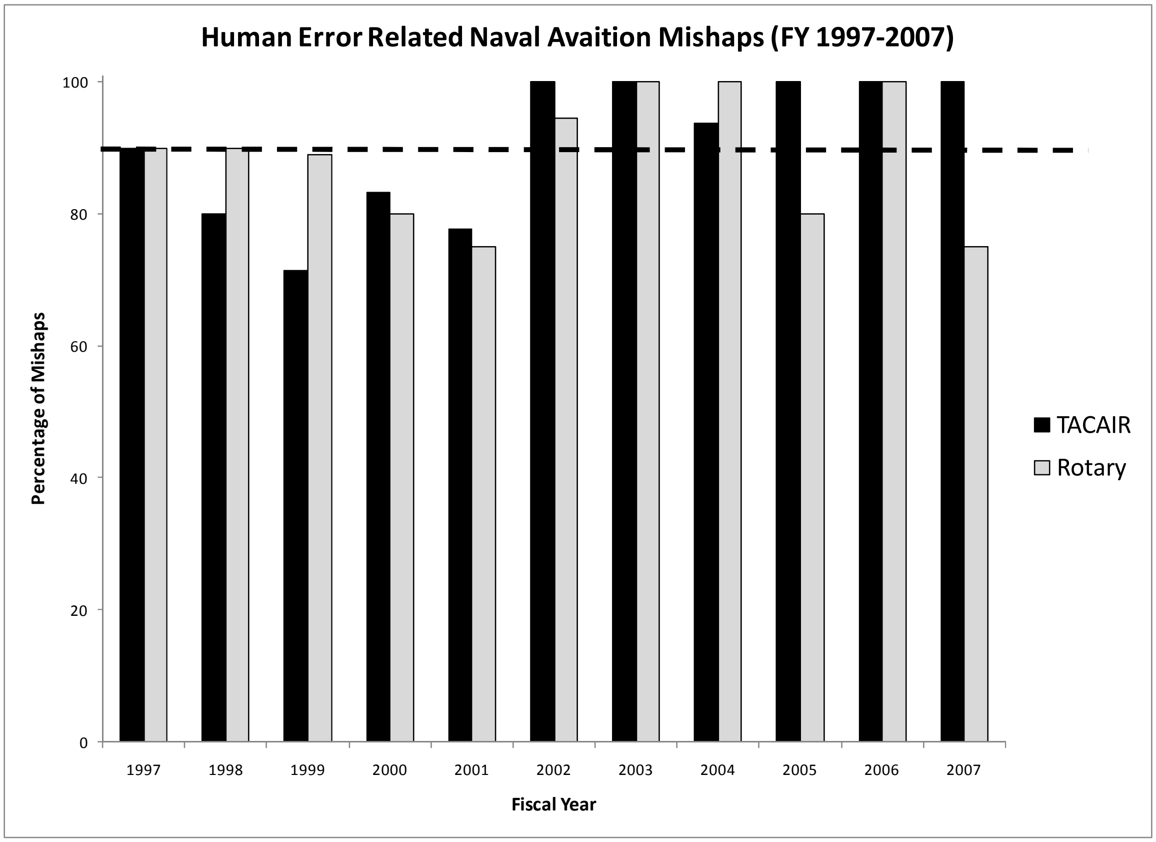

Figure 13. Percent of mishaps (for TACAIR and rotary wing aircraft) involving human error. The solid line indicates the average mishap percentage for both airframes.

3. Mishaps Due to CRM Failures

A total of 68.9% of the 238 mishaps were determined to have involved at least one form of CRM failure that was determined to have had a causal factor in the mishap. The TACAIR community showed that 62.9% of its mishaps were related to CRM failure(s). Rotary wing aircraft percentages were, generally, higher than TACAIR at an average of 77.6% of mishaps (see Figure 14).

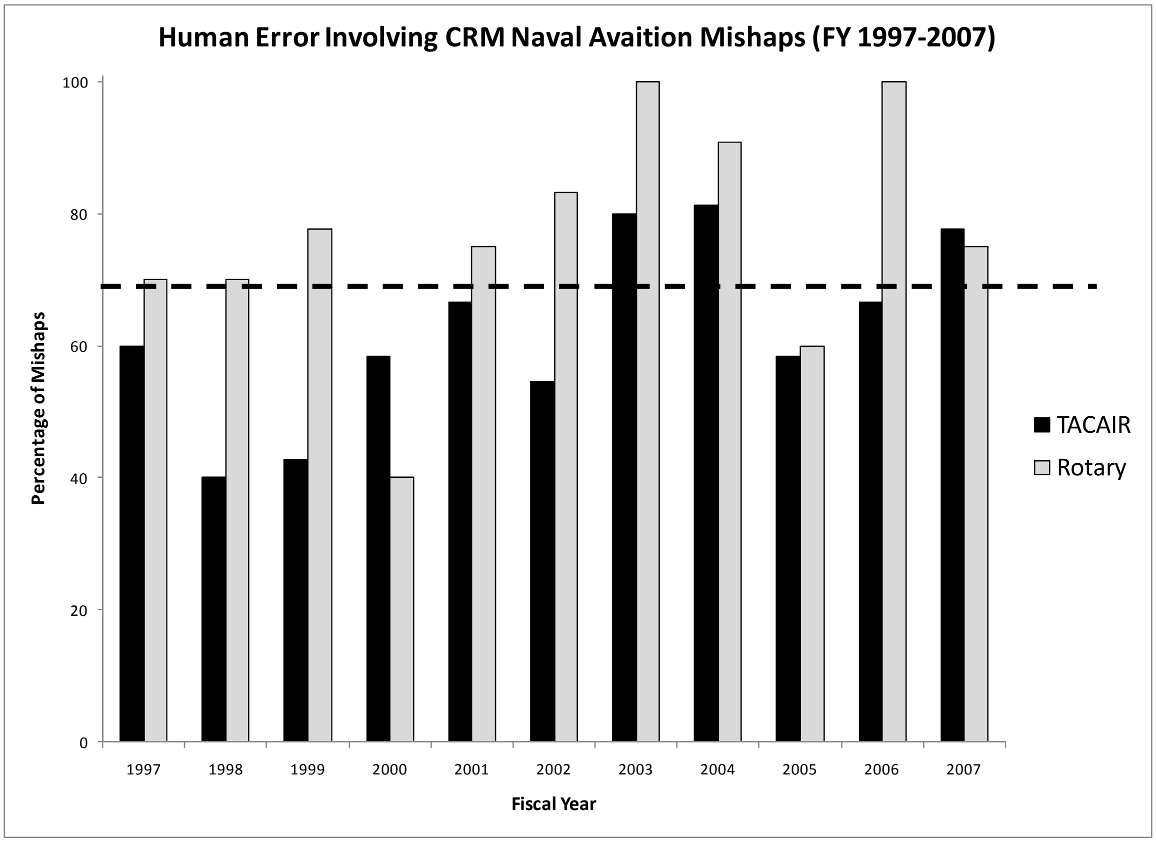

Figure 14. Percent of mishaps (for TACAIR and rotary wing aircraft) involving at least one form of CRM failure. The solid line indicates the average mishap percentage for both airframes.

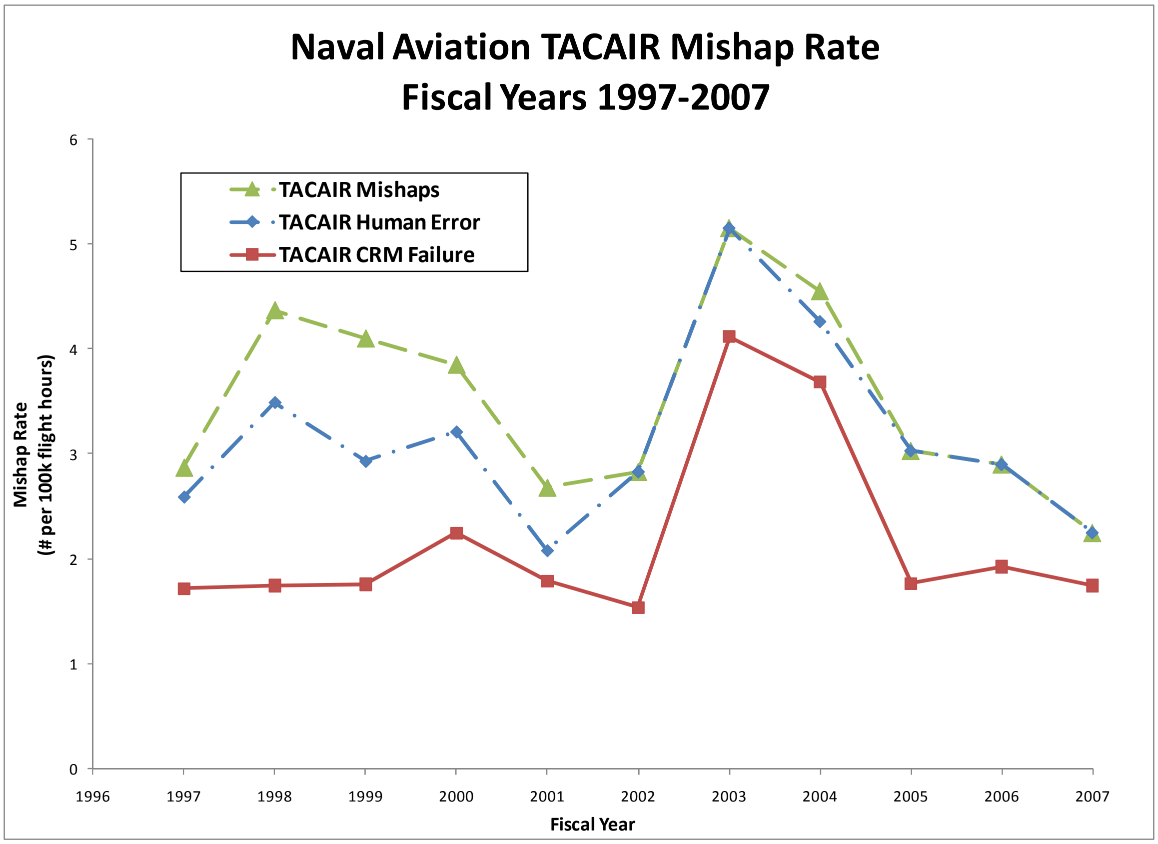

Figure 15. Mishap rates for TACAIR from fiscal year 1997 to 2007. The overall TACAIR mishap rate is indicated by triangles with a dashed line. Mishap rates involving some form of human error are represented by diamonds with a dashed/dotted line. Mishap rates attributed to at least one CRM failure are represented by squares with a solid line.

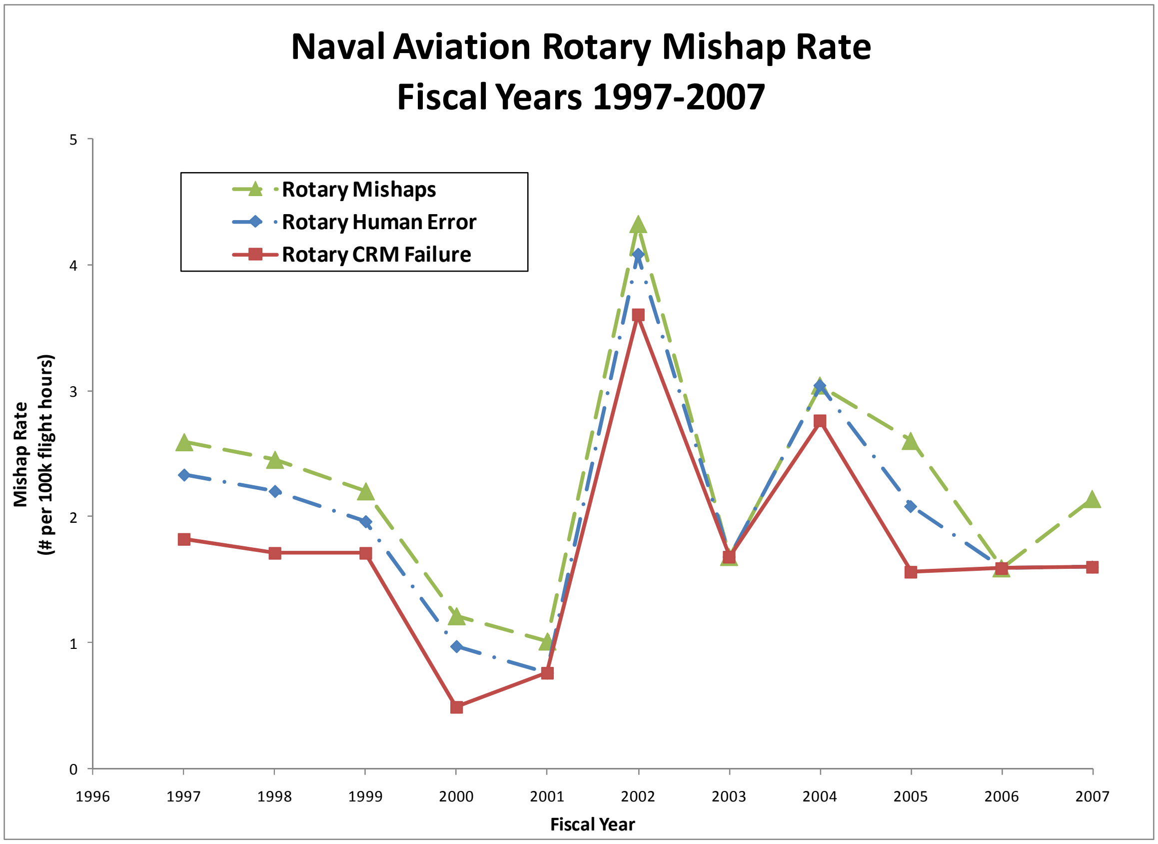

Figure 16. Mishap rates for rotary wing aircraft from fiscal year 1997 to 2007. The overall Rotary wing mishap rate is indicated by triangles with a dashed line. Mishap rates involving some form of human error are represented by diamonds with a dashed/dotted line. Mishap rates attributed to at least one CRM failure are represented by squares with a solid line.

D. ANALYSIS AND DISCUSSION

Class A mishap rates in naval aviation have remained at lower levels despite the assorted fluctuations seen in the safety data from fiscal year 1997 to 2007. Even over the past decade, the average mishap rate is slightly lower (2.85 per 100,000 flight hours) as compared to the mishap rate (3.84 per 100,000 flight hours) found by Wiegmann and Shappell (1999) in their study from 1990 to 1996. In addition, both TACAIR and rotary wing aircraft average mishap rates have faintly decreased compared to the prior study. TACAIR had dropped from an average rate of 4.38 in 1990-96 to 3.48 across 1997-2007. Similarly, rotary wing aircraft decreased from 2.98 to 2.26 over the same time frame. TACAIR continues to exhibit a higher mishap rate than rotary wing aircraft as was observed in Wiegmann and Shappell's (1999) review. Considerations of tactical jet versus rotary wing aircraft mission sets were ascribed as the reasons for this difference in earlier research (Wiegmann & Shappell, 1999). The hazardous nature of TACAIR maneuvers and training could account for these differences. Finally, consideration must be given to the operational tempo seen as a result of wartime evolutions over the time frame examined in this study. This may be a positive indication that with more mission flights vice training flights, the mishap rate has remained fairly constant.

Despite a minimal decrease in overall mishap rates in naval aviation from the early 1990s to the early twenty-first century, the percent of mishaps attributed to human error has not followed suit. Human error accounted for nearly 78% of naval aviation mishaps as compared to 90% seen over the past decade. This obvious increase in human error related mishap percentages suggest human factor issues continue to present problems in naval aviation. However, the increase in percentage related to human error could be a result of acceptance and awareness of the contribution of human error in aviation. As the organization becomes more attentive to potential human-related errors during operations, these causal factors are more readily identifiable. Another explanation for the increase in percentage may be that mishap rates due to mechanical or environmental reasons have significantly decreased to insignificant levels, perhaps due to technological advances and safety prevention programs. This trend would result in a larger percentage of mishaps caused by a human-related failure.

The average percent of aviation mishaps over the past several years involving at least one occurrence of failure in CRM is approaching 70%. These results fail to provide sufficient support for the proposed hypothesis that CRM has had a positive impact on the mishap rate over the past decade. Wiegmann and Shappell (1999) also found that 56% of all aircrew-related mishaps involved a failure of CRM. Similarly, rotary wing aircraft mishaps associated with CRM failures were consistently higher than TACAIR mishaps (77.6% versus 62.9%). This disparity was not as sizeable as found in previous research which found a difference of 33.9% higher percentage of rotary wing aircraft. Rotary wing percentages have decreased while TACAIR percentages have increased. Again, CRM concepts and awareness may result in a larger percentage of TACAIR mishaps being coded as CRM-related. As for the higher rotary wing percentage of mishaps due to CRM failures, helicopter crews are traditionally larger and often more complex than seen in TACAIR aircraft. The mishaps which occur in routine operations of rotary airframes are likely to have a high probability of being crew-related. Compared to the results found in the early 1990s, the percentage has dropped, but remains higher than desired.

E. SUMMARY

CRM failures continue to plague naval aviation and contribute to a high percentage of mishaps. However, mishap rates in naval aviation have decreased slightly over the past decade, but not to insignificant levels ascribed to mechanical and environmental factors. It may be true that human error is unavoidable and can never be eliminated as long as humans play an integral role in aviation systems. CRM programs seek to reduce the number of mishaps attributed to poor crew communication and coordination. As seen from this and previous CRM research, assessing the direct impact of training on aviation mishaps and safety is difficult. CRM might be the key component in maintaining low mishap rates in naval aviation, despite the large percentage of mishaps attributed to human error. There may not be much left to attribute as a cause of the mishap. CRM has, most likely, helped in avoiding several accidents saving money and lives. However, it is nearly impossible to account for the success rate in terms of mishap prevention associated with CRM training.

Nevertheless, examining the effectiveness of training in terms of its impact on the organization is important. Although no direct conclusions can be made regarding the success of CRM in terms of deterrence of aviation mishaps, this study provided valuable and updated information regarding naval aviation mishap rates over the past decade. Evaluating a program in terms of results alone is clearly not sufficient (Kirkpatrick, 1996). This is the primary reason this research effort attempted to look at the value of naval CRM training across different levels of Kirkpatrick’s (1976) hierarchy including the reactions, learning (attitudes and knowledge), and organizational impact.

THIS PAGE INTENTIONALLY LEFT BLANK

VII. CONCLUSIONS AND RECOMMENDATIONS

A. SUMMARY

This thesis applied a multi-faceted evaluation technique to provide an objective measure of the effectiveness of the Navy and Marine Corps CRM program. Using Kirkpatrick's (1976) hierarchy, four studies pertaining to the reactions, attitudes, knowledge, and outcomes of Naval CRM were examined. Reactions of naval aviators were found to be generally favorable toward the utility of CRM concepts taught in each aviation command. No specific trends were found other than a slightly positive increase in aviator reaction responses following fiscal years 2002 and 2003. Positive reactions hint at the success of CRM instruction and delivery methods throughout the Navy and Marine Corps. Favorable reactions were consistent with prior CRM research indicating a continued acquiescence for the training. While trainee satisfaction does not ensure effective training has occurred, contentment in the program suggests learners may be more motivated to participate and acknowledge the value of instruction.

Learning in the Naval CRM program was also examined through both an attitudinal and knowledge-based assessment. No differences were found between aviators in their attitudes toward general concepts addressed in CRM training. Adjustments to the delivery method and instruction of techniques addressing specific CRM concepts warrant further consideration. For example, increasing awareness of other crewmembers' stress levels for tactical jet aviators deserves contemplation for future CRM development. The responses to the NAHF attitude questionnaire suggest senior aviators tend to be more receptive, than expected, to subordinates expressing a difference in opinion. In CRM training, emphasizing the fact senior officers indicated their willingness to listen to junior aviators who may be reluctant to speak up in CRM training could prove beneficial. Because this research relied upon a modified version of the CMAQ to assess naval aviator attitudes, CRM concepts heavily rooted in cognitive processes were not addressed. Therefore, the NAHF questionnaire could be supplemented by additional research methods to investigate the posture of CRM concepts related to decision making, situational awareness, and workload management (O'Connor et al., 2008).

A knowledge assessment of Naval CRM (KANCRM) found no significant differences between aviators based upon rank or airframe. This finding may suggest CRM training, despite the variation in instruction from command to command, is fairly consistent in emphasizing the non-technical concepts within the program. However, the observed low mean knowledge scores may be a warning sign that certain concepts are not being addressed adequately in the program. Knowledge assessments are too often excluded from CRM effectiveness evaluations. The KANCRM, although limited in scope, provides a cursory assessment technique to help quantity a fraction of the learning as a result of CRM training. Expansion of the assessment might present additional metrics to further investigate how well specific CRM concepts are retained. Furthermore, routinely administering a knowledge evaluation may be considered.

The third level of assessment, behaviors, was omitted from this study due to limited time, feasibility and resources. However, it is important to take into consideration the impact a training program may have on behaviors. Behavioral changes are key indicators in identifying whether knowledge acquired in training truly transfers to improved or altered behaviors on the job. Evaluation of behaviors can be witnessed through observational studies as well as simulated environments. Several previous studies evaluating behavior changes have suggested CRM successfully changes trainee behaviors (O'Connor et al., 2002; Salas et al., 2001) while other studies provided arguments of unconvincing behavioral changes (Salas et al., 2006). More research is recommended to study the effects of learning on behavioral changes in the Naval CRM program.

Ultimately the vitality of any training program relies on a combination of favorable reactions, increased learning, and effective behavioral changes. But often the measure of a program's outcome will compel an organization to implement or continue to fund the training. This is the reason most training evaluations only estimate the return on investment in terms of bottom line results. Safety and mission accomplishment remain the top priorities in naval aviation. An examination of naval aviation was performed using safety data from fiscal years 1997 through 2007 and found that mishap rates have remained relatively constant (under five mishaps per 100,000 flight hours). These mishap rates are similar to what has been observed since CRM was first applied in naval aviation in the early 1990s (Yacavone, 1993; Wiegmann & Shappell, 1999). The direct impact CRM training has on mishap rates and preventing incidents remains unclear. On the other hand, mishaps are infused with human factors issues. CRM may be the chief contributing factor that already keeps mishap rates at such low levels, and the remaining percent of mishaps cannot be eliminated as human error will always exist.

B. CRM HSI CONSIDERATIONS

Crew interaction, pilot experience, flow of information, team cognition, and fatigue are all aspects of human factors that need to be further explored in aviation to maximize safety and performance. From a Human Systems Integration (HSI) perspective, CRM concepts are intended to coordinate all available resources of a system, from both human and machine, to ensure the goal of improving performance and safety is accomplished. Today's aviators are required to interact routinely with computerized systems increasing the workload and synchronization complexity of managing both the crew and machine. HSI and human performance considerations multiply and become more significant as the equipment used by high reliability organizations becomes more sophisticated and complex.

The research in this thesis described several interdisciplinary tradeoffs between cognitive processing, team coordination, and personnel safety seen in U.S. Navy and Marine Corps aviation. Techniques taught in CRM training must be considered in the design of a system, and especially in a system of systems. Maurino (1999) suggests "effective CRM training requires dropping the piecemeal strategies largely favored in the past in favor of a system approach, because implementing CRM does not simply mean training pilots, controllers, and mechanics, but developing the organization" (p. 230). Manpower, personnel, training and education (MPT&E) are critical inputs required to infuse CRM concepts into routine flight operations. Employing sound human factors engineering principles into aviation system acquisitions will reinforce good CRM practices. Incorporating HSI into the future of aircraft design and CRM development is necessary to make the most of the Navy and Marine Corps human capital. In fact, the U.S. Department of Defense (DoD) now mandates that program managers of new military systems must consider HSI in the design process. “The Program Manager shall have a plan for HSI in place early in the acquisition process to optimize total system performance, minimize total ownership costs, and ensure that the system is built to accommodate the characteristics of the user population that will operate, maintain, and support the system” (Department of Defense, 2008: 60).

C. RECOMMENDATIONS AND FURTHER RESEARCH

Technological advances continue to attempt to improve reliability and safety in aviation. Design considerations throughout the acquisition process must also continue to expound on integrating humans, including our cognitive and physiological capabilities, as fundamental aspects of aviation systems. Nevertheless, CRM training continues to be the primary method in mitigating the number of mishaps due to human crew coordination error in aviation. The program must go beyond simple training requirements and become more in grained into the culture. Perhaps the evidence found in this study suggests a transformation of cultural acceptance is gradually occurring.

Issues mentioned in this thesis suggest supplementary research and evaluation of the Navy’s CRM program be considered. Specifically, it is important to gather additional metrics on the reactions, learning, behavior and outcomes as a result of naval CRM training. Collecting CRM feedback could easily be accomplished through a routine and standardized evaluation. Establishing a CRM baseline assessment, using non-aviation participants, could provide added benefit in understanding the effectiveness of the training. Data collected from research can then be analyzed in longitudinal studies to assess the impact of CRM on human factor behaviors. It is also recommended for naval aviation to consider developing or modifying specific training objectives for each type of airframe. Designing training that meets the idiosyncrasies of TACAIR, rotary wing, and other fixed wing aircraft might increase the usefulness of the training. Accomplishing this task may require separate evaluations of the CRM course material for each airframe within a command.

CRM programs have begun to extend beyond aviation into other high reliability organizations (Flin, O’Connor, & Mearns, 2002). Lessons and assessments from other

CRM programs can be applied to make improvements to the Navy and Marine Corps aviation program. Additionally, results from multi-level evaluations, as seen here, may aid in designing, implementing, or assessing future CRM team coordination training. The future of CRM in U.S. Naval Aviation must be considered. Making CRM an adaptive and responsive training program allows training to be tailored to the ever-changing operating environment. Rather than waiting years or decades to analyze outcomes of CRM training, regular evaluations provide critical information and feedback to quickly revise curriculum to meet mission requirements. The U.S. Army's Aircrew Training Coordination – Enhanced (ACT-E) is an example of a program that seeks to continuously adapt to changes in aviation requirements.

Finally, as aviation technology matures and the use of Unmanned Aerial Vehicles (UAV) becomes ubiquitous, CRM training also must adapt its concepts to account for such changes. Additional research in human and automated machine interaction play a vital role in future CRM concept development, identifying needed capabilities and non-technical skills required for autonomous aviation operators.

D. CONCLUSION

Extensive funding devoted to improving safety, both through procedures and technology advances in aviation, have greatly reduced the mishap rate over the past several decades. However, one fact remains: human error is a principal contributor toward mishaps in naval aviation. CRM training remains a fundamental technique used to ensure effective team coordination in aviation with the goal to minimize preventable accidents. Therefore, it is important to continue to assess whether CRM is meeting its goal in the Navy of "improving mission effectiveness by minimizing crew preventable errors, maximizing crew coordination, and optimizing risk management" (Chief of Naval Operations, 2001).

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A. NAHF QUESTIONNAIRE

Please answer the following items by using the following scale in writing your response beside each item:

A Disagree Strongly	**B** Disagree Slightly	**C** Neutral	**D** Agree Slightly	**E** Agree Strongly

____1. Even when fatigued, I perform effectively during critical times in a flight.

____2. I let other members of the aircrew know when my workload is becoming (or is about to become) excessive.

____3. My decision-making ability is as good in emergencies as in routine flying operations.

____4. I am more likely to make judgment errors in an emergency.

____5. A debriefing after each flight is an important part of developing and maintaining effective teamworking.

____6. In abnormal situations, I rely on my superiors to tell me what to do.

____7. I am less effective when stressed or fatigued.

____8. My performance is not adversely affected by working with an inexperienced or less capable member of the aircrew.

____9. Successful flightdeck management is primarily a function of the Aircraft Commander's flying proficiency.

____10. If I perceive a problem with the flight, I will speak up, regardless of who might be affected.

____11. The pre-flight brief is important for safety and effective teamworking.

____12. Human factors <u>councils</u> are regarded as non-punitive in my squadron.

____13. Aircrew should monitor each other for signs of stress or fatigue.

____14. Personal problems can adversely affect my performance.

____15. Truly professional aircrew can leave personal problems behind when flying.

____16. Aircrew should mention their stress or physical problems to other team members before or during a flight.

____17. Good communication and crew coordination are as important as technical proficiency for flight safety.

A Disagree Strongly	B Disagree Slightly	C Neutral	D Agree Slightly	E Agree Strongly

____18. Effective crew coordination requires aircrew to consider the personal work styles of other team members

____19. Aircrew should alert others to their actual, or potential, work overload.

____20. A true professional does not make mistakes.

____21. Human factors <u>boards</u> are regarded as non-punitive in my squadron.

____22. NATOPS procedures are always followed in my squadron.

____23. The specific roles and responsibilities of aircrew in an emergency are identified during the preflight brief.

____24. Written procedures are necessary for all in-flight situations.

____25. Uncertain situations often require quick decision making.

____26. If the human factors council is about to discuss one of the council members, then that individual always leaves the room.

____27. Aircrew should be aware of, and sensitive to, the personal problems of other team members.

____28. Junior aircrew should not question the Aircraft Commander's decisions in emergencies.

____29. It is better to agree with other crew members than to voice a different opinion.

____30. Aircraft Commanders who encourage suggestions from aircrew are weak leaders.

____31. The Aircraft Commander should take physical control and fly the aircraft in emergency and non-standard situations.

____32. The pilot flying the aircraft should verbalize plans for procedures or maneuvers and should be sure that the information is understood and acknowledged by other crew members.

____33. Aircrew should not question actions of the Aircraft Commander, except when they threaten the safety of the flight.

____34. Crew members share responsibility for prioritizing activities in high workload situations.

____35. There are no circumstances (except total incapacitation) where the co-pilot should assume command of the aircraft.

____36. Junior aircrew should not question the Aircraft Commander's decisions during normal operations

____37. How frequently in your squadron are junior personnel afraid to express disagreement with more senior personnel (please circle below)?

A. Very frequently
B. Frequently
C. Sometimes
D. Seldom
E. Very seldom

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B. ANALYSIS OF THE NAHF QUESTIONNAIRE

Skewness and Kurtosis of the NAHF:

NAHF analysis

		N Valid	Skewness	Std. Error of Skewness		Kurtosis	Std. Error of Kurtosis			Missing
My stress	1Fatigued effective ®	410	-0.37	0.12	-3.06	-1.02	0.24	-3.1	-4.2	1
Others stress	2Let others know workload	410	-0.84	0.12	-6.96	0.00	0.24	-7.0	0.0	1
My stress	3decision emergency routine ®	411	-0.46	0.12	-3.82	-0.90	0.24	-3.8	-3.7	0
My stress	4Judgement errors emergency	410	-0.15	0.12	-1.28	-1.14	0.24	-1.3	-4.8	1
Comms	**5Debrief is important**	**411**	**-3.63**	**0.12**	-30.15	**14.95**	**0.24**	**-30.1**	**62.2**	**0**
Command	6Abnormal, rely on superiors	411	0.68	0.12	5.66	-0.49	0.24	5.7	-2.0	0
My stress	7Less effective stressed fatigued	410	-0.80	0.12	-6.63	-0.27	0.24	-6.6	-1.1	1
My stress	8Performance not affected by inexperience	410	0.22	0.12	1.86	-1.19	0.24	1.9	-5.0	1
Command	9Successful flightdeck AC proficiency ®	402	0.10	0.12	0.78	-1.10	0.24	0.8	-4.5	9
Comms	**10If perceive problem speak up**	**411**	**-2.81**	**0.12**	-23.36	**9.30**	**0.24**	**-23.4**	**38.7**	**0**
Comms	**11Pre-flight important**	**411**	**-6.08**	**0.12**	-50.48	**43.73**	**0.24**	**-50.5**	**182.0**	**0**
Not CRM	12HF councils nonpunitive	404	-1.11	0.12	-9.17	0.39	0.24	-9.2	1.6	7
Others stress	**13Airecrew should monitor**	**411**	**-2.94**	**0.12**	-24.41	**10.09**	**0.24**	**-24.4**	**42.0**	**0**
My stress	14Personal problems affect performance	411	-0.96	0.12	-7.94	0.15	0.24	-7.9	0.6	0
Rules	15Truly professional behind ®	411	-0.27	0.12	-2.23	-1.14	0.24	-2.2	-4.8	0
Others stress	16Airecrew mention stress fatigue	407	-1.48	0.12	-12.25	2.58	0.24	-12.3	10.7	4
Comms	**17Good comms important as technical**	**411**	**-3.49**	**0.12**	-29.02	**15.55**	**0.24**	**-29.0**	**64.7**	**0**
Others stress	18Effective coordination personal work styles	411	-1.60	0.12	-13.29	3.16	0.24	-13.3	13.2	0
Others stress	19Alert others to mistakes	411	-1.68	0.12	-13.98	5.25	0.24	-14.0	21.9	0
Not CRM	**20True professional no mistakes**	**411**	**3.63**	**0.12**	30.16	**14.31**	**0.24**	**30.2**	**59.6**	**0**
Not CRM	21HF boards nonpunitive	402	0.00	0.12	0.02	-0.85	0.24	0.0	-3.5	9
Rules	22NATOPS followed	409	-0.68	0.12	-5.65	-0.50	0.24	-5.6	-2.1	2
Comms	23Roles and respons emergencies id	411	-1.58	0.12	-13.13	2.97	0.24	-13.1	12.4	0
Rules	24Written procedures	411	0.83	0.12	6.92	-0.31	0.24	6.9	-1.3	0
Not CRM	25Uncertain situations quick	411	-0.81	0.12	-6.75	-0.44	0.24	-6.8	-1.8	0
Not CRM	26HF council, then leave	400	-0.71	0.12	-5.86	-0.64	0.24	-5.9	-2.6	11
Others stress	27Aircrew aware and sensitive	411	-1.03	0.12	-8.55	1.01	0.24	-8.5	4.2	0
Command	28Junior not question in emergency	400	1.81	0.12	14.85	3.61	0.24	14.8	14.8	11
Rules	**29Better to agree**	**400**	**2.66**	**0.12**	21.78	**7.64**	**0.24**	**21.8**	**31.4**	**11**
Command	**30AC suggestions are weak ®**	**400**	**4.80**	**0.12**	39.38	**25.37**	**0.24**	**39.4**	**104.2**	**11**
Command	31AC control in emergency	389	0.32	0.12	2.60	-0.82	0.25	2.6	-3.3	22
Comms	**32Pilot flying verbalize**	**399**	**-2.44**	**0.12**	-19.93	**6.89**	**0.24**	**-19.9**	**28.3**	**12**
Command	33Aircrew not question AC	399	1.22	0.12	9.98	0.88	0.24	10.0	3.6	12
Command	**34Crew members responsible for priority**	**398**	**-2.06**	**0.12**	-16.86	**4.32**	**0.24**	**-16.9**	**17.7**	**13**
Command	35No circumstances when co-pilot	389	1.23	0.12	9.97	0.51	0.25	10.0	2.1	22
Command	36Junior question during normal	399	1.36	0.12	11.15	1.55	0.24	11.2	6.4	12
Not CRM	37Junior disagree	402	0.01	0.12	0.04	-0.06	0.24	0.0	-0.2	9

Count	Subscale	Items	Alpha (before)	Dropped	Alpha (after)
6	My stress	*1,3* ,4,7,8,14	0.09	3,4,7,14	0.52
6	Stress of others	2,13,16,18,19,27,	0.57		0.57
6	Comms	5,10,11,17,23,32	0.51		0.51
9	Command	6,*9* ,28,*30* ,31,33,34,35,36	0.03	6,28,31,33,35,36	0.54
4	Rules	*15* ,22,24,29	0.0007	Drop scale	

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C. KNOWLEDGE ASSESSMENT FOR NAVAL CRM (KANCRM)

Please answer the following questions based on your experience and training:

What percentage of accidents in your community is attributed to human error?

- 0-10%
- 20-30%
- 50-60%
- 80-90%

What is the most common reason skilled operators fail to obtain good situation awareness in a high workload situation?

- Data is unavailable
- Difficulty detecting data
- Memory loss
- Failure of scan

You are working in a potentially high-risk situation with a more senior person. Identify whether each of the statements below is PASSIVE, which is ASSERTIVE, and which is AGGRESSIVE.

	PASSIVE	ASSERTIVE	AGGRESSIVE
"Sir/Ma'am. I think that perhaps this may not be the best thing to do."			
"Sir/Ma'am. There may be a better way to do this operation. In my opinion we should do ...."			
"Sir/Ma'am. There is no way we should do this."			

During operations, something unexpected happens. There is no procedure available. You have seconds to make a decision to attempt to prevent a bad outcome, what is the best thing to do?

- Consider all of the possible options, and select the best.
- React the best you can to the situation based upon your experience.
- Do nothing.

During a time limited, non-standard, operational situation what is the most effective team communication strategy?

- The team should talk a lot about what is happening in an attempt to solve a problem.
- Use the minimum amount of communication necessary in an attempt to solve the problem.
- Don't talk. This distracts from thinking about the problem.

During a non-standard situation, which statement below best describes what inexperienced individuals tend to do?

- Forget the procedures.
- Misjudge the time available, and react too quickly.
- Take too much time to think about the issue.
- Panic

The following are all characteristics of an effective pre-mission brief EXCEPT:

- ◯ Assigning of roles and responsibilities
- ◯ Rapid information dissemination
- ◯ Professional
- ◯ Interactive

What are some of the advantages of procedural, or rule-based decision making? (identify all that apply)

- ☐ You do not need to be an expert.
- ☐ You do not need to understand the purpose of every step.
- ☐ Produces solution for unfamiliar problem
- ☐ Useful when trying to solve a novel problem.

The normal need for sleep to maximize performance in a 24 hour period is:

◯ 1-3 hours ◯ 4-6 hours ◯ 7-9 hours ◯ 10-12 hours

Which of the following statement is true regarding the relationship between stress and performance:

- ◯ Performance is optimized when an individual is experiencing no or a very little amount of stress.
- ◯ Performance is optimized when an individual is experiencing a moderate amount of stress.
- ◯ Performance is optimized when an individual is experiencing an excessive amount of stress.
- ◯ Stress has no effect on an individual's performance.

LIST OF REFERENCES

Alkov, R. A. (2005). *The Naval Safety Center's Aircrew Coordination Training Program.* Retrieved April 24, 2009, from Neil Krey's CRM Developers Forum: http://www.crm-devel.org/resources/paper/alkov.htm

Alkov, R. A. (1989). The U.S. Naval aircrew coordination program. *Proceedings of the Fifth International Symposium on Aviation Psychology* (pp. 368-371). OH: Ohio State University.

Alliger, G. M., Tannenbaum, S. I., Bennet, W., & Holly Traver, A. S. (1997). A Meta-Analysis of the Relations Among Training Criteria. *Personel Psychology* (50), 341-358.

Alliger, J., & Janak, E. (1989). Kirkpatrick's levels of training criteria: Thirty years later. *Personnel Psychology, 42*, 331-341.

Baker, D., Bauman, M., & Zalesny, M. D. (1991). Development of aircrew coordination exercises to facilitate training transfer. (R. Jensen, Ed.) *Proceedings of the 6th International Symposium on Aviation Psychology*, 314-319.

Booher, H. (2003). *Handbook of human systems integration.* New Jersey: John Wiley & Sons.

Brannick, M., Prince, A., Prince, C., & Salas, E. (1995). The measurement of team process. *Human Factors, 37*, 641-651.

Brown, D. (1987). Introduction to MAC CRM training. In H. Orlady, & H. Foushee, *Cockpit resource management training: Proceedings of the NASA/MAC workshop (NASA CP-2455)* (pp. 132-134). Moffet Field, CA: NASA-Ames Research Center.

Chief of Naval Operations. (2001). Crew Resource Management Program. OPNAVIST 1542.7C. Washington, D.C.

Cooper, G., White, M., & Lauber, J. (1980). Resource management on the flightdeck. *Proceedings of a NASA/Industry workshop (NASA Conference Proceedings-2120).* Moffett Field, CA: NASA-Ames Research Cente.

Darby, J. (2008). Course evaluations: a tendency to respond “favourably” on scales? *Quality Assurance in Education, 16* (1), 7-18.

Davies, J. (2001). Medical Application of Crew Resource Management. In E. Salas, C. Bowers, & E. Edens, *Improving teamwork in organizations.* Mahwah, NJ: Lawrence Erlbaum Associates.

Department of Defense (2008). Operation of the Defense Acquisition System, DoD Instruction 5000.02.Washington D.C.: Office of the Secretary of Defense.

Department of the Navy. (2007). *Naval Aviation Safety Program: OPNAVINST 3750.6R.* Washington, D.C.: Chief of Naval Operations.

DeVellis, R. (1991). *Scale development: Theory and applications.* London: SAGE Publications.

Endsley, M. (1995). Toward a theroy of situation awareness in dynamic situations. *Human Factors, 37*, 32-64.

Federal Aviation Administration (FAA). (1989). *Cockpit resource management training (Advisory Circular 120-51A).* Washington, D.C.: U.S. Department of Transportation.

Federal Aviation Administration. (2004). *Advisory Circular No 120-51E: Crew resource management training.* Washington, D.C.

Flin, R. (1996). *Sitting in the hot seat: Leaders and teams for critical incident management.* Chichester, UK: John Wiley & Sons.

Flin, R., O'Connor, P., & Crichton, M. (2008). *Saftey at the Sharp End: A guide to non-technical skills.* Burlington, VT: Ashgate.

Flin, R., O'Connor, P., & Mearns, K. (2002). Crew Resource Management: Improving safety in high reliability industries. *Team Performance Management, 8*, 68-78.

Foushee, H., & Helmreich, R. (1988). Group interaction and flightcrew performance. In E. Wiener, & D. Nagel, *Human Factors in Aviation* (pp. 189-227). San Diego, CA: Academic Press.

Fowlkes, J., Lane, N., Salas, E., Franz, T., & Osler, R. (1994). Improving the measurement of team performance: The TARGETs methodology. *Military Psychology, 6* (1), 47-61.

Gliner, J., Haber, E., & Weise, J. (1997). Use of controlled vignettes in evaluation: Does type of response method make a difference? *Evaluation and Program Planning, 22*, 313-322.

Goldstein, I., & Ford, J. (2002). *Training in Organizations: Needs Assessment, Development, and Evaluation* (4th ed.). Belmont, CA: Wadsworth.

Gregorich, S., & Wilhelm, J. (1993). Crew resource management training assessment. In E. L. Wiener, B. G. Kanki, & R. L. Helmreich, *Cockpit Resource Management* (pp. 173-196). San Diego: Academic Press.

Gregorich, S., Helmreich, R., & Wilhelm, J. (1990). The structure of cockpit management attitudes. *Journal of Applied Psychology, 75* (6), 682-690.

Guzzo, R., & Dickson, M. (1996). Teams in organizations: recent research on performance and effectiveness. *Annual Review of Psychology, 47*, 307-338.

Hartel, C., Smith, K., & Prince, C. (1991). Defining Aircrew Coordination: Searching mishaps for meaning. *Sixth International Symposium on Aviation Psychology*. Columbus: Ohio State University.

Helmreich, R. (1984). Cockpit management attitudes. *Human Factors, 26*, 583-589.

Helmreich, R., & Foushee, H. (1993). Why Crew Resource Management? Empirical and theoretical bases of human factors training in aviation. In E. L. Wiener, B. G. Kanki, & R. L. Helmreich, *Cockpit Resource Management* (pp. 3-45). San Diego, CA: Academic Press Inc.

Helmreich, R., & Merritt, A. (1998). *Culture at Work in Aviation and Medicine: National, Organizational and Professional Influences.* Aldershot: Ashgate.

Helmreich, R., & Wilhelm, J. (1991). Outcomes of Crew Resource Management Training. *The International Journal of Aviation Psychology, 1* (4), 287-300.

Helmreich, R., Foushee, H., Benson, R., & Russini, W. (1986). Cockpit resource management: exploring the attitude-performance linkage. *Aviation, Space, and Environmental Medicine, 57*, 1198-1200.

Helmreich, R., Merritt, A., & Wilhelm, J. (1999). The Evolution of Crew Resource Management Training in Commercial Aviation. *International Journal of Aviation Psychology, 9* (1), 19-32.

Hofstede, G. (1980). *HofstCulture's Consequences: International differences in work-related values.* Beverly Hills, CA: Sage.

Holt, R., Boehm-Davis, D., & Beaubien, J. (2001). Evaluating Resource Management. In E. Salas, C. Bowers, & E. Edens, *Improving teamwork in organizations* (pp. 165-188). Mahwah, NJ: Lawrence Erlbaum Associates.

Irwin, C. (1991). The impact of initial and recurrent Cockpit Resource Management training on attitudes. *In Proceedings of the Sixth International Symposium on Aviation Psychology* (pp. 344-349). Columbus: Ohio State University.

Jentsch, F., & Smith-Jentsch, K. (2001). Assertiveness and Team Performance: More than "Just Say No". In E. Salas, C. Bowers, & E. Edens, *Improving Teamwork in Organizations.* Mahwah, NJ: Lawrence Erlbaum.

Karp, R., Condit, D., & Nullmeyer, R. (1999). Cockpit/Crew Resource Management for Single-Seat Fighter Pilots. *In Proceedings of the Interservice/Industry Training, Simulation and Education Conference* [CD-ROM]. Arlington, VA: National Training Systems Assoc.

Katz, L. C., & Grubb, G. N. (2003). *Enhancing U.S. Army Aircrew Coordination Training.* Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Kerlinger, F. (1996). *Foundations of behavioural research.* New York: Holt, Rinehart & Winston.

Kirkpatrick, D. L. (1996). Evaluation. In R. L. Craig, *The ASTD Training and development Handbook* (4th ed., pp. 294-312). New York, NY: McGraw-Hill.

Kirkpatrick, D. L. (1976). Evaluation of training. In R. L. Craig, & L. R. Bitte, *Training and development handbook* (pp. 18.1-18.27). New York, NY: McGraw Hill.

Kirkpatrick, D. L. (1998). *Evaluation training programs.* San Francisco: Berrett-Koehler.

Klein, G., Orasanu, J., Calderwood, R., & Zsambok, C. (1993). *Decision Making in Action: Models and Methods.* Norwood, NJ: Ablex Publishing Co.

Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcome to new method of training evaluation. *Journal of Applied Psychology, 78* (2), 311-328.

Lauber, J. K. (1987). Cockpit resource management: Background studies and rationale. In H. Orlady, & H. Foushee, *Cockpit resource management training: Proceedings of the NASA/MAC workshop (NASA CP-2455)* (pp. 5-14). Moffett Field, CA: NASA-Ames Research Center.

Leedom, D. K. (1990). *Aircrew coordination training and evaluation for Army rotary wing aircrews: Summary of research for fiscal year 1990.* Fort Rucker: U. S. Army Research Institute Aviation Research and Development Activity.

Lovell, K. (2007). Getting the value out of evaluation. *Training Journal*, 43-46.

Mathieu, J., Tannenbaum, S., & Salas, E. (1992). Influences of Individual and Situational Characteristics on Measures of Training Effectiveness. *The Academy of Management Journal, 35* (4), 828-847.

Maurino, D. (1999). Crew resource management: A time for reflection. In D. Garland, J. Wise, & V. Hopkin, *Handbook of Human Factors* (pp. 215-234). Mahwah, NJ: Lawrence Erlbaum Associates.

Naval Aviation Schools Command. (2008, May 7). *Seven Skills of CRM*. Retrieved January 21, 2009, from https://www.netc.navy.mil/nascweb/crm/standmat/seven_skills/MA.htm

Naval Safety Center. (2006). *Aviation 3750.* Norfolk, VA: Naval Safety Center.

Naval Safety Center. (2009). *Aviation Safety Survey Program*. Retrieved March 23, 2009, from http://www.safetycenter.navy.mil/aviation/surveys.htm

Njaa, O. (2000). On the use of accident scenarios. *Paper presented at the SPE conference on Health, Safety, and the Environment in Oil and Gas Exploration.* Stavanger, Norway.

O'Connor, P. (2007). The nontechnical causes of diving accidents: Can U.S. Navy Divers learn from other industries? *Undersea Hyperb Med, 34* (1), 51-59.

O'Connor, P., Flin, R., & Fletcher, G. (2002). Methods used to evaluate the effectiveness of CRM training: A literature review. *Journal of Human Factors and Aerospace Safety, 2* (3), 217-234.

O'Connor, P., Campbell, J., Newon, J., Melton, J., Salas, E., & Wilson, K. (2008). Crew Resource Management Training Effectiveness: A Meta-Analysis and Some Critical Needs. *International Journal of Aviation Psycholog , 18* (4), 353-368.

O'Connor, P., Hahn, R. G., & Nullmeyer, R. (in press). The Military Perspective. In: R.L. Helmreich, B. G. Kanki, & J. Anca (Eds.). *Crew Resource Management*. San Diego. CA: Elsevier

O'Connor, P., Hahn, R., & Salas, E. (in press). The U.S. Navy's Crew Resource Management Program: The past, present, and recommendations for the future. In P. O'Connor, R. Hahn, E. Salas, P. O'Connor, & J. Cohn (Eds.), *Human performance enhancements in high-risk environments: insights developments, and future directions from military research.* Shreveport, LA: Paragon Press.

O'Connor, P., O'Dea, A., & Melton, J. (2007). A methodology for identifiying the human factors causes of U.S. Navy diving accidents. *Human Factors, 49*(2), 214-226.

Orasanu, J. (1990). *Shared mental models and crew decision making (CSL Report No. 46).* Princeton University, Cognitive Science Laboratory, Princeton, NJ.

Orasanu, J. (1993). Decision-making in the cockpit. In E. Wiener, B. Kanki, & R. Helmreich, *Cockpit Resource Management* (pp. 137-168). San Diego, CA: Academic Press.

Orasanu, J., & Salas, E. (1993). Team decision making in complex environments. In G. Klein, J. Orasanu, R. Calderwood, & C. Zsambok, *Decision Making in Action.* Norwood, NJ: Ablex.

Orlady, H. W., & Foushee, H. C. (1987). *Cockpit Resource Management Training* (Vols. (NASA CP-2455)). Moffett Field, CA: NASA-Ames Research Center.

Oser, R., Salas, E., Merket, D., & Bowers, C. (2001). Applying resource management training in naval avaition: A methodology and lessons learned. In E. Salas, C. Bowers, & E. Edens, *Improving teamwork in organizations* (pp. 283-301). Mahwah, NJ: Lawrence Erlbaum Associates.

Phillips, J. (1991). *Handbook of Training Education and Measurement* (2nd ed.). Houston, TX: Gulf publishing.

Phillips, J. (1996). Measuring the Results of Training. In R. Craig, *The ASTD Training and Development Handbook* (pp. 313-341). New York, NY: McGraw-Hill.

Prince, C., & Salas, E. (1989). Aircrew Performance: Coordination and skill development. In D. Daniel, E. Salas, & D. Kotick, *Independent research and independent exploratory development (IR/IED) programs: Annual report for fiscal year 1988* (pp. 35-40). Orlando, FL: Naval Training Systems Center.

Prince, C., & Salas, E. (1993). Training and research for teamwork in the military aircrew. In E. Wiener, B. Kanki, & R. Helreich, *Cockpit Resource Management* (pp. 337-365). San Diego, CA: Academic Press Inc.

Reason, J. (1990). *Human Error.* Cambridge: Cambridge University Press.

Roscoe, S., & Williges, R. (1975). Measurement of Transfer of Training. In S. Roscoe, *Aviation Psychology* (pp. 182-193). Ames, IA: The Iowa State University Press.

Salas, E., & Klein, G. (2001). *Linking Expertise and Naturalisitc Decision Making.* Mahwah, NJ: Lawrence Erlbaum.

Salas, E., Burke, C., Bowers, C., & Wilson, K. (2001). Team training in the skies: Does crew resource management (CRM) training work? *Human Factors*, 641.

Salas, E., Fowlkes, J. E., Stout, R. J., Milanovich, D. M., & Prince, C. (1999). Does CRM training improve teamwork skills in the cockpit?: Two evaluation studies. *Human Factors*, *41*, 326-343.

Salas, E., Rhodenizer, L., & Bowers, C. (2000). The Design and Delivery of Crew Resource Management Training: Exploiting Available Resources. *Human Factors, 42* (3), 490-511.

Salas, E., Wilson, K., Burke, C., & Wightman, D. (2006). Does CRM training work? An update, extension and some critical needs. *Human Factors*, *14*, 392-412.

Schiewe, A. (1995). On the acceptance of CRM-methods by pilots: Results of a clusteranalysis. (R. Jensen, Ed.) *Proceedings of the Eighths International Symposium on Aviation Psychology*.

Secretary of the Air Force. (1995). *Cockpit/Crew Resource Management Training Program.* Washington, D.C.: U.S. Government Printing Office.

Seva, R., Gutierrez, A., Duh, H., & Chong, J. (2007). An Evaluation of CRM Attitudes of Filipino Pilots in Four Philippine Aviation Companies. *International Journal of Aviation Psychology, 17* (3), 285-298.

Shappell, S., & Wiegmann, D. (2000). *The Human Factors Analysis and Classification System—HFACS.* Washignton, D.C.: Office of Aviation Medicine; Federal Aviation Administration.

Shappell, S., & Wiegmann, D. (1996). U.S. Naval Aviation Mishaps, 1977-92: Differences Between Single- and Dual-Piloted Aircraft. *Aviation, Space, and Environmental Medicine, 67* (1), 65-69.

Stout, R. J., Salas, E., & Fowlkes, E. J. (1997). Enhancing teamwork in complex environments through team training. *Group Dynamics: Theory, research, and practice, 1*, 169-182.

Stout, R. J., Salas, E., & Kraiger, K. (1996). The role of trainee knowledge structures in aviation psychology. *The International Journal of Aviation Psychology*, *7*, 235-250.

Thordsen, M., Klein, G., & Wolf, S. (1990). Observing team coordination within Army rotary-wing aircraft crews (KATR90C2-90-05-Z). Yellow Springs, OH: Klein Associates.

Wickens, C., & Hollands, J. (2000). *Engineering Psychology and Human Performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Wiegmann, D., & Shappell, S. (1999). Human error and crew resource management failures in naval aviation mishaps: A review of U.S. Naval Safety Center data, 1990-96. *Aviation, Space, Environmental Medicine, 70* (12), 1147-1151.

Wiener, E., Kanki, B., & Helreich, R. (1993). *Cockpit Resource Management.* San Diego, CA: Academic Press Inc.

Yacavone, D. (1993). Mishap trends and cause factors in naval aviation: a review of Naval Safety Center data, 1990-1996. *Aviation, Space, Environmental Medicine, 64* (5), 392-395.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Paul E. O'Connor
   Naval Postgraduate School
   Monterey, California

4. Michael E. McCauley
   Naval Postgraduate School
   Monterey, California

5. John Owen
   NAVAIR
   Orlando, Florida

6. M. Quinn Kennedy
   Naval Postgraduate School
   Monterey, California

7. CAPT Brad Robinson
   Naval Safety School
   Pensacola, Florida

8. CAPT Ed Langford
   Naval Safety Center
   Norfolk, Virginia