2002

Aircrew Adaptive Decision Making: A Cross-Case Analysis

Constance Ann Gillan EdD
University of San Diego

University of San Diego
School of Education
Leadership Studies Program

Aircrew Adaptive Decision Making: A Cross-Case Analysis

In Partial Fulfillment
of the Requirements for the Degree

Doctor of Education

by

Constance Ann Gillan

Mary Woods Scherr, Ph D., Chair
Robert Donmoyer, Ph D., Member
Robert Nullmeyer, Ph D., Member

ABSTRACT OF THE DISSERTATION

Aircrew Adaptive Decision Making: A Cross-Case Analysis

Although the accident rate for military aviation has declined significantly from earlier decades, during the 1990's it reached a plateau. Human error in the cockpit still accounts for over 80% of the aircraft mishaps resulting in loss of life or over one million dollars in damage. Decision error has been a contributing factor for approximately 60% of these mishaps. The purpose of this research was to investigate aircrew process performance variables as predictors of decision-making outcomes.

This study was modeled on elements of previous research in naturalistic decision making. Data were collected for cross-case analysis of the role experience plays in efficient decision strategy selection and use in an uncertain, dynamic high stakes environment. Multiple raters evaluated eight novice and eight experienced military aircrews at seven decision points in a 20-minute flight scenario conducted in a full motion flight simulator. Other raters independently rank ordered the quality of the final outcome. A comprehensive approach to collecting and analyzing data included: (1) development and use of a behaviorally-anchored assessment instrument, (2) use of a digitally integrated presentation of audio/video and flight data, and (3) development of context-specific analytical frameworks and models of observed behaviors and metacognitive processes. Results included inferential and descriptive statistics of process/outcome scores, instructor comments, excerpts of cockpit recordings, participant interviews, and field notes.

The study findings were: (1) high individual and collective crew experience had a significant positive effect on process and outcome scores, (2) there was no statistically reliable difference in process scores between experience levels in the three procedurally-based events, (3) experienced crews performed better than novice crews in the four less structured events, (4) novice crews' process/ outcome correlation did not approach significance, (5) a strong positive correlation of process/outcome scores was found for experienced crews in the two most challenging (i.e., unstructured) scenario events, (6) qualitative analysis revealed strong relationships between performance and crew interactions/attributes, and (7) in dynamic, time critical situations, the use of adaptive decision-making strategies led to better performance outcomes.

DEDICATION

This dissertation is dedicated to my parents
Edward and Marian O'Rourke

ACKNOWLEDGMENTS

I will be forever indebted to my dissertation committee chair and advisor, Professor Mary Woods Scherr, for her wisdom, enthusiasm, and vital support she provided throughout my tenure in the USD Leadership Program. I am also grateful for the inspiring recommendations, guidance and enthusiasm of my committee members, Professor Bob Donmoyer and Bob Nullmeyer, Ph. D.. The encouragement, focus, and well-grounded perspective these role models provided will undoubtedly guide my future efforts within the Human Factors arena.

This study would not have been possible without the help and assistance of the dedicated and talented members of United States Naval aviation community. A special thank you to Sea Control Wing Commodores Langley, Buckner, Hepfer and Kelly. Additionally, Commander Greg LaBuda, Squadron Commanding Officers, and the participating aircrews from the active duty and Reserve S-3 community were key contributors to this work. I would also like to acknowledge the importance of the Naval Aviation Human Factors Quality Management Board and the Naval Safety Center for defining the vision and goals that guided the study; to reduce the Naval aviation aircraft mishap rate while improving combat readiness and operational effectiveness.

I am grateful for the quantitative analysis support provided by the United States Air Force Warfighting Research Laboratory in Mesa, Arizona commanded by Major General Paul Nielsen and directed by James Brinkly and Colonel Gerald Straw. The contributions of Peter Crane, Ph. D. and Bill Stock, Ph. D. were instrumental in completing the quantitative analysis of the study data.

A very special thank you to my qualitative analyst-in-chief, Sheldon Hunt, who was instrumental in integrating the study findings into pragmatic constructs, models and schemas. Sheldon exceeded every expectation for finding the truth and providing the insight that a former S-3 pilot "Instructor of the Year" brought to the discussion and interpretation of the research findings. Similarly, the assistance of LCDR Paul Tanks, an S-3 Naval Flight Officer, was instrumental in designing a relevant and uniquely challenging scenario that provided the framework for studying adaptive decision making of S-3 aircrews.

This dissertation study has been a truly collaborative effort. I would like to acknowledge the following individuals whose involvement was instrumental in the completion of this effort. The contract simulator instructors at VS-41 provided their invaluable expertise to probe for metacognitive processes during the crew debriefs. A sincere thank you to Mr. Frank Cavanaugh, Mr. Charlie Beauchesne, Mr. Mike Lettieri, and LCDR Paul Tanks who made themselves available on short notice scheduling. Mr. Ken Fallis provided technical support with computer data files while Jim Blanchard, Sc. D created human factors analysis profiles for novice crews. Alan Spiker, Ph.D., Gail Pamisano, Ed.D., CDR Steve Lilley, and Ms. Robin Elz assisted with the review and the revision of the manuscript and provided timely, critical comments on successive drafts of the work. Mr. Chuck Wightman and Mr. Merle Vogel provided invaluable technical support and assistance. A most sincere and heartfelt thanks to all of you.

TABLE OF CONTENTS

LIST OF APPENDICES

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

CHAPTER ONE

STATEMENT OF THE ISSUE

Introduction

Time critical, adaptive decision making used to rapidly assess non-routine dynamic situations underpins success in the dynamic aviation environment. This study explored decision-making approaches of Navy faced with recognizing and controlling unplanned and emergency situations within seconds or minutes given conflicts and gaps in critical information. Research was conducted to investigate the use of adaptive decision-making approaches by novice and experienced aircrews in a specific simulator context representative of their decision-making domain.

The decision-making process explored lies on a decision making continuum between the time consuming, linear, analytical approach at one extreme, and a reactive, procedurally-based approach on the other end. The focal point on this continuum for this study is an adaptation to these decision strategies. Somewhere in the center of this continuum is an intuitive adaptive decision making approach that synthesizes domain experience with insightful knowledge for use in atypical contexts that require immediate problem resolution.

In order to better understand the relationship of the use of adaptive decision strategies and aircrew performance under stressful conditions data were collected in a simulated flight event along with the aircrews' self-reported cognitive processes in the

event debrief. Data analyses were performed to determine if: (a) previous "like-kind" or transferable experiences in the aircraft are critical to how aircrews remember and report their interpretations of vague and concurrent cues, (b) aircrew experience levels determine the differences in thought patterns of aircrews, and (c) a correlation exists between aircrew situational assessment, decision-making processes and related decision process outcomes.

Background

In 1996, the U.S. Navy's most senior aviators chartered a Human Factors Quality Management Board to discover best practices worldwide and conceive interventions to forestall human error from the proximate (cockpit) level to the resource (organizational) level. After a year of military and commercial aviation program reviews of safety and operational readiness strategies, several initiatives were adopted. One of these initiatives by the Naval Safety Center created a Human Factors Analysis Classification System (HFACS) utilizing Reason's Model of Human Error (Reason, 1990). This mishap causal factor taxonomy was created and populated with discrete behaviors and conditions proven to cause aircraft mishaps. Using the HFACS taxonomy, an analysis was conducted by the of 110 (81 Tactical Air/ 29 Helicopter) Navy/Marine Corps Class "A" (those involving one million dollars or more and loss of life) aviation mishaps that occurred from 1990 through 1996.

The results of this analysis of tactical fixed wing aircraft Class "A" mishaps found that 63 % of the mishaps included decision errors. These decision errors were categorized to include one or multiple decision errors per mishap as follows: (a) wrong response in an emergency (16 mishaps), (b) improper use of flight controls (12 mishaps),

(c) poor decision (7 mishaps), (d) exceeded pilot/aircrew ability (5 mishaps), (e) misdiagnosed emergency (5 mishaps), (f) misinterpreted/misused instruments (5 mishaps), (g) inappropriate maneuver (4 mishaps), (h) failed to recognize extremis (3 mishaps), (i) improper approach/landing (3 mishaps), (j) improper takeoff (1 mishap), and (k) used incorrect data (1 mishap).

Meeting the goal of eliminating human error mishaps demands the adoption of more effective instructional strategies, along with better organization climate, increased supervision and accountability across the span of naval aviation aircraft platforms and functional areas. The need for improved shore-based and shipboard-training capabilities has become increasingly critical as the Navy increases its operational commitments with fewer, less experienced personnel. This need created a triad of powerful training initiatives from the operational level for training aircrews: (a) use advanced technology and software applications to systematically capture and measure aircrew performance, (b) inclusion of human factors and performance measurement and analysis in instructor training, and (c) employment of advanced curriculum development methodologies to integrate critical thinking and problem solving applications throughout the aviation training continuum.

Statement of the Issue

While the military aviation accident rate has declined significantly from earlier decades, it has reached a plateau since 1991. Human error still accounts for over 80% of incidences in both military and civilian flight mishaps. United States Naval Safety Center analysis of military aircrew accidents in 1998 revealed insignificant change over time in the percent of mishaps attributed at least in part to decision-making errors and

lapses of judgment by aircrews. In 1998, aircrew decision-making errors contributed to almost 60% of all Class "A" mishaps. Since current policy and levels of technology restrict data collection during actual flight environments (i.e., most military aircraft are not equipped with cockpit voice recorders or "black boxes" nor do they always operate in environments observed by radar which makes for comparatively easy flight reconstruction) it is difficult to observe aviators under stressful conditions of uncertainty in their natural decision making environments. Modeling the most proficient decision makers in a particular domain requires access to their cognitive and behavioral data. The best available setting to this researcher for collecting aircrew performance data and providing decision-making experiences to aircrews is currently in the full-motion flight simulator. To date, there has been little progress in the approach to improve the decision-making capabilities of aviators beyond assessment of overt decision making capabilities. Much of the research used to support aviation training has been done largely with novices in non-flight environments (e.g., Payne, Bettman & Johnson, 1988; Prince, Hartel, & Salas, 1992; Stokes, Kemper & Kite, 1997).

Several problem areas impede decision making data collection in realistic aviation environments for practical as well as theoretical use. First, most research in decision making thought processing patterns related to probabilities and outcomes has been highly structured using content impoverished stimuli (e.g., use of gambling simulations by Payne et al., 1988). Second, Bowers, Jentsch, and Salas (2000) point out that there is a misconception that that a universal skill set is appropriate for all aviation platforms and operational contexts. Third, without the use of sophisticated data collection and analysis tools it is difficult to collect and replicate detailed flight data to

support metacognitive recall of processes reflecting considerations and actions by aircrews in time-critical processes. Fourth, limitations in obtaining time and access to experts in realistic work environments have made field research with any experts, much less aviators, difficult (Shanteau, 1988).

In aviation contexts, investigations must begin to include methodologies that produce data that capture how aviators adapt their decision-making processes as they gain flight experience to effectively manage uncertainty and risk in the cockpit. Shanteau (1986) found that expert (i.e., more experienced) decision-makers make decisions differently than non-experts (i.e., novices). More experienced aviators are more likely to have an experiential knowledge base that is used to identify exceptions to rules, to assess situations quickly and define how much time to allot to problem solving. They take more calculated risks to manage outcomes. These differences are important; they open an area of research that can better definition of training requirements for aviators at all levels of experience.

Previous practical investigation of experienced aircrews' implicit decision making processes as proposed in this study has been limited to cognitive task interviews. Designed to elicit decision-making thought protocols based on past experience and "what if?" paper-based scenarios, these analyses have been extremely valuable; however, they lack the real-time, detailed data that can be observed and collected in a flight simulator or actual flight event (e.g., verbal protocols, information-acquisition behavior, task shedding, and response times). Since content influences psychological processes (Tetlock, 1985), an actual flight event in a high fidelity aircraft simulator would produce a more realistic context for the study of collective expert and novice aircrew decision

making especially under conditions in which aircrews, as teams, face uncertain conditions.

Study Purpose

The purpose of study was to collect observational and metacognitive data from aircrews to identify adaptations to decision making strategies as articulated by better performing crews. Furthermore, data was sought that would illuminate the relationship between aircrew characteristics and aircrew functioning processes relative to multi-crewed cockpit performance.

This study addressed the utility of multi-method data collection and analysis to capture aircrew decision-making considerations and flexibility in assessing and planning approaches to resolve flight related problems in time critical situations. Technical and crew coordination data were collected to measure aircrew processes; latency and quality of decision outcome data were studied to discover possible relationships between these study variables. Consequently, results were reviewed to determine and clarify essential skills, procedures, and strategies required to train more effective aircrews facing similar situations.

Hypotheses and Research Questions

This study proposed three hypotheses and three research questions that form a starting point of the cumulative process to explore and identify variables in an environmentally valid situation (i.e., a realistic scenario representative of a natural situation). In this context, potentially relevant variables were observed and considered by both the research team and the participants. The study hypotheses were tested using inferential statistics to explore possible relationships between the performance process

variables and outcome variables and aircrew experience level. Combined analysis and interpretation of quantitative and qualitative data served as the basis for a possible integrative theoretical and training approach that may have relevance beyond the one naturalistic situation represented by the study scenario.

The research hypotheses and questions related to a single flight scenario were:

Hypothesis 1: Experienced aircrews will receive better process ratings (i.e., technical skill, workflow, information sharing, consensus building, operational risk assessment and management and back-up routines) than novices in handling a specific in-flight emergency situation involving uncertainty.

Hypothesis 2: Decisions (i.e., outcome rankings) made by experienced aircrews in an uncertain situation involving a specific in-flight emergency are rated as being of higher quality than those made by novice aircrews.

Hypothesis 3: Aircrew process ratings are correlated with the decision outcome ratings for a specific in-flight emergency involving uncertainty.

Research Question 1: What adaptive recognitional/metacognitive decision-making strategies emerge from aircrews in a specific in-flight emergency situation involving uncertainty?

Research Question 2: In what way do adaptive/metacognitive decision-making patterns differ among successful and less than successful aircrews challenged with a specific in-flight emergency involving uncertainty?

Research Question 3: What characteristics/factors seem to define the most successful aircrew outcomes in a specific in-flight emergency involving uncertainty?

This research study investigated whether experienced military aircrews from the fleet with at least 700 hours of flight experience in a particular aircraft had different metacognitive approaches to identify, assess and resolve airborne problems than less experienced crews in the same aircraft. The 700-flight hour level is the total amount of flight time accumulated by an S-3 aviator with at least one year in a fleet squadron. The hypotheses aimed at investigating possible relationships between experience level and various aspects of specific aircrew behaviors and skills that might correlate with approaches and biases in judgments and decision-making. Aspects of aircrew performance (e.g., technical skills, procedural skills, inter-personal skills, higher level cognitive skills, crew coordination, etc.) within and among the two different experience levels studied were rated by expert observers and analyzed for statistically significant correlation with final mission outcome. Qualitative analysis methodology was used to create situational awareness and decision taxonomies representing aircrew actions and thought processes during critical decision points in the scenario.

Methodology

This field-based comparative case study of aircrew decision making under conditions of uncertainty was conducted at a United States naval aviation training facility in San Diego. Aircrews flying the S-3B model aircraft in routine training syllabi at the basic (i.e., novice) and advanced (i.e., experienced) levels of naval aviation carrier-based training were asked to volunteer as study participants.

A case study approach was designed and adopted to explore the relationship between aircrew situational awareness and decision-making processes and the outcomes

of experienced and novice aircrews. A flight scenario in a full-motion simulator was used to trigger two dangerous event conditions in an ambiguous, information-limited context that gave the aircrews little to immediate response time. Unexpected changes in the flight and landing environment were followed with greater instability and novelty to assess whether aircrews would change their information gathering/use and decision-making strategies. Data were analyzed to determine if there was a meaningful or statistically significant correlation between aircrew flight experience and the successful handling of sequentially more problematical emergency situations.

Study data consisted of quantitative performance assessments during the realistic 20-minute simulator scenario followed by participant interviews related to crew decision processes and outcomes. Process variables were comprised of technical and crew resource management tasks; decision outcome data consisted of aircraft configuration and position at the end of the scenario. Up to five subject matter experts that collected observational data used a behaviorally-anchored Likert scale instrument designed for the study. Evaluation criteria reflected both general team process and event-specific standards of performance. These subject matter experts served as independent observers/assessors. They were senior instructors at the training command for the model aircraft flown by the participant crews.

Immediately following the 20-minute scenario, semi-structured, recorded debrief/interviews between 40 and 60 minutes in duration were conducted with aircrews to collect data related to the research questions. To enhance recall of cues and decisions in the scenario for both participants and independent observers, a stimulated recall strategy was employed during the post-scenario debrief/interview. A computerized

digital replay of selected instruments in the aircraft instrument panel, aircraft aspect, as well as audio and video of the aircrew over every minute of the scenario event was available and used during portions of the debrief/interview to initiate recall of technical data as well as the participants' reactions to events and their situational awareness and decision making during the study scenario.

Conceptual Framework

Over the last 15 years, studies involving complex decision-making style adaptations in response to uncertain, dynamic conditions with various populations in real-world domains have contributed to the movement away from the classic "rational" model of decision making. There is now a new focus on the type of decision making built on an individual's intuitive recognition and resolution of atypical situations using his experience base. This recognitional approach to decision making is known as naturalistic decision making (Klein, 1993, 1998; Zsambock, 1997). The decision making process in uncertain situations can be generally categorized by assigning it to one of three groups with strategy shifts occurring with the number of options (Payne, 1982) and time constraints: (a) procedural or rule-based (e.g., use of standard operating procedures, checklists, etc.) typically used to for automatic, rapid responses, (b) analytical which consumes much time in an attempt to gather all available information in various combinations and assigns a fixed set of possible meanings to cues/information to weight all possible outcomes, and (c) adaptive or creative problem solving which is a process that attempts to deal with limitations in information gathering and processing in time critical situations.

In adaptive problem solving the first available action that satisfies the immediate need for a desired outcome or "end game" is selected to meet the situations time constraints. Adaptive decision-making processes often synthesize intuition, experience, and a tailored application of the normal procedures and/or techniques to obtain the desired objective. Early indications in work using the Recognition-Primed Decision (RPD) Model (Klein, Calderwood, & Clinton-Cirocco, 1986; Klein, 1989, 1993) has provided evidence through deep interview techniques and field observations that there are differences in the way experienced and inexperienced people working as teams recognize and respond to problems in high stress environments. In general, research has found that experienced persons rely on their long-term memory and make time-critical, high stakes decisions in potentially volatile situations by using sophisticated strategies based on recognition and reactions to trends and patterns from their experiences. Inexperienced people, by contrast, tend to rely on various recalled preferences and ignore or prematurely discount alternative hypotheses or options. An extension of the Klein et al (1986). RPD model is the Recognition/Metacognition (R/M) model that describes a set of critical thinking strategies to verify results of recognition and problem correction in novel situations through metacognitive processes (Cohen, Adelman, Tolcott, Bresnick, & Marvin, 1993; Cohen, Freeman, & Wolf, 1996). This model was developed using critical incident interviews with U.S. Army infantry personnel and U.S. Navy shipboard personnel. Both the RPD and R/M models of naturalistic decision making along with Zsambok's (1997) Aviation Decision Process Model, which specifies cognitive processes involved in aviation decision making, were of primary importance in conceptually framing this research.

Definition of Terms

For the purposes of this study, variables were operationally defined as follows:

Adaptive Decision Making

Adaptive decision making involves a shift in decision-making strategy from procedural or analytical approaches to meet the demands of a time-critical, threatening, dynamic situation with an undefined structure. This shift in decision strategy is utilized because procedural or analytic approaches are too time consuming in a real-time, dynamic situation. The requirement to accelerate the decision making process is typically brought about by a need to trade accuracy for speed, shifting decision criteria, and task shedding.

Naturalistic Decision Making

This type of decision making is concerned with how individuals use their knowledge to make decisions in the face of unruly problems embedded in dynamic task contests. Eight factors characterize decision making in naturalistic settings: ill-structured problems, uncertain dynamic environments, shifting, ill defined or competing goals, action/feedback loops, time stress, high stakes, multiple players, and organizational goals and norms. It is not likely that all eight factors will be at their most difficult levels in any one situation or setting but often several of these factors will complicate the decision task (Orasanu & Connolly, 1993).

Creative problem solving

This process requires the use of critical thinking skills when no response is readily available as a standard operating procedure, there is no guidance for dealing with the malfunction in the aircraft operating manuals, and when the crew has not been trained to assess or manage the situation. " In these cases they must invent a candidate solution

to meet their goals and evaluate its adequacy in the light of existing constraints" (Orasanu, 1997, p.54).

Metacognition

This term bridges the areas between (a) decision making and memory, (b) learning and motivation, and (c) learning and cognitive development. It describes our "knowledge about how we perceive, remember, think, and act - - that is, what we know about what we know" (Metcalfe & Shimamura, 1996, p. xi). It is the ability of a person to reflect on his or her conscious awareness of judgments and experiences. It is a supplemental recognition process used to verify and improve recognition of a situation.

Meta-recognition

Meta-recognitional skills probe for flaws in recognized assessments and plans, try to patch up any weaknesses found, and evaluate the results. Meta-recognitional processes include: (a) identification of evidence-conclusion relationships (or arguments) within the evolving situation or plan, (b) processes of critiquing that identifying problem in the arguments that support the situation model or plan which can result in the discovery of problems of incompleteness, unreliability, or conflict, (c) processes of correcting that respond to these problems, and (d) a control process called *quick test*, which regulates critiquing and correcting. The quick test considers the available time, the costs of error, and the degree of uncertainty or novelty in the situation (Cohen & Freeman, 1996).

Limitations of the Study

The findings of this study are limited to identification, description, and analysis of participant aircrews' decision-making processes and outcomes in a single case

scenario. This study did not attempt to predict the success of various decision-making styles found to be used in the study scenario to other situations or domains. The case study approach was used to focus on the patterns of aircrew experiences in a particular naturalistic situation. Each case may provide data to expand the cognitive schema but there is no attempt to generalize findings to other flight situations or populations. Additionally, because the study participants may not be representative of the S-3B aircraft population there is also no attempt to generalize study findings to the S-3B aircraft aircrew population. However, the relevancy of the findings may be useful to enhance specific cognitive skill training approaches and naturalistic decision making theories.

Significance of the Study

Results of the study are useful to practical aircrew and instructor training applications. The larger issues addressed in this research are those of when and how adaptive decision making employment has been successful in meeting immediate operational needs in an unstable environment. Study data collected via combined qualitative and quantitative approaches resulted in gaining more details of various aircrew problem recognition and problem solving approaches than either method employed separately. Advanced levels of description that emerged from the study provide details that will assist training practitioners in extending the depth of process and performance feedback to aircrews and instructors. Evidence presented serves as a summary delineating aircrews' behavior and functioning used to assimilate, process and investigate various information (cues) with particular attention being paid to methods for dealing with uncertainty. Practical aircrew and instructor training applications could be

designed to provide models and discussion points for training in operational best practices to include risk assessment and management.

Significant advances in aircrew training require domain-specific information that is used to predict, explain, and understand aircrew decision-making biases in their problem solving strategies under conditions of uncertainty and limited response time. Recent U.S. Air Force aircrew studies (Spiker, Tourville, Silverman & Nullmeyer, 1996; Spiker, Silverman, Tourville & Nullmeyer, 1998; Nullmeyer & Spiker, 1999) have shown that there is variability in performance processes and outcomes among even the most experienced aircrews. The data analysis from the present study provides more data on experienced/novice decision-making processes and also extends the analysis to provide additional metacognitive and inferential data on several variables not targeted in similar combinations in other studies with military aviators. The findings of this study may shed more light on why some aircrews succeed in a given situation while others do not. Additionally, the study research design, methodology and results could be useful in complementing other specific measures of performance, measures of effectiveness research, and training applications that have only used explicit data. Finally, interpretations from analyzed data captured in this domain-specific case study could be constructive in providing generalizations that contribute to individual and team behavioral decision theory. Results of the study were also used as feedback to the training command and squadrons.

Summary

Current administrative and training approaches do not routinely address aircrew deficiencies in decision making. Improvements in training critical thinking skills will most certainly lead to improvements in mission performance, saved equipment, and ultimately, saved lives. Investigations that capture effective cognitive approaches, thought protocols, and performance behaviors of the most successful aircrews in time-critical decision making situations require combined data collection and analysis methodologies. This research study integrated three naturalistic decision making models (i.e., Recognition-Primed Decision (RPD) model, the Recognition/Metacognition (R/M) model, and the Aviation Decision Process model) that allowed an in-depth examination of a single situation across and between novice and more experienced aircrews. This framework assisted in the discovery of rich, thematic connections between aircrew experience levels and decision processes and outcomes.

CHAPTER TWO

REVIEW OF THE LITERATURE

Introduction

Despite training military personnel in analytical and procedural approaches to decision making, these strategies do not consistently result in optimal performance under varying dynamic conditions with variable degrees of complexity and stress. Emerging analyses from recent research studies (Pounds & Fallesen, 1995; Pascual & Henderson, 1997; Zsambok, 1997; Bainbridge, 1999) have found that military personnel are using a recognitional approach to find solutions when they have had an experience with a similar problem type. Although other decision-making strategies are commonly used, the recognitional, or adaptive, naturalistic decision-making style has been found to dominate in dynamic environments. Findings from these decision-making studies have determined that expert decision-makers naturally leaned toward using experienced-based approaches rather than time consuming, concurrent comparisons of options. Using the perspective that risky decision making is highly sensitive to variations in the task environment, Payne (1985) suggests that an approach for studying risky choice involves examining several stages of information processing behavior that are highly contingent upon task and context variables.

Cognitive approaches to decision making that focus on information processing have become a more dominant force over normative decision theory in recent years as reported by Maule (1985). The focus for much of the research has been, and continues to be, along two paths. First, an information processing focus to identify the separate stages of information processing and explain how each operates. Second, a bigger picture focus on "complex cognitive skills like problem solving and considering how stages operate and interact in the execution of these skills" (p. 62).

Humphreys and Berkeley (1985) point out that one area that has received little attention is the motivation behind preferred decision-making strategies of individuals within a team that may affect team performance. They suggest it is important to investigate "where uncertainty enters into the process of conceptualizing a decision problem, thus enabling discussions on how it is handled by the decision maker at each point" (p.258). In real life decision problems, individuals conceptualize decision problems and then gage their ability to act upon an uncertain situation. To manage the inherent uncertainty, individuals select from coping mechanisms such as trial and error, use of feedback, and treating uncertain parts of the decision problem as certain to attempt to manage a resolution.

Theoretical Framework

Within the literature examined, four processes underpin the theoretical framework of the research: adaptive decision-making strategies, situational awareness, and effects of experience on aircrew performance, and team performance and evaluation. These four areas will be discussed as they relate to the research proposed to examine the influence of experience in situational assessment and decision making with experienced

and novice aircrews in a situation involving uncertainty, high stakes, and time constraints in a representative naturalistic setting.

Naturalistic Decision Making (NDM) is the study of how people use their experience to make decisions in field settings (Zsambok, 1997) and this approach is "often informed by practical problems as opposed to testing hypotheses derived from theories" (p. 12). There are several features that distinguish the descriptive theories associated with naturalistic decision making from analytical or procedural approaches to problem solving or decision making under conditions of risk. In time-critical situations where information is uncertain, a simplified task-shedding strategy that includes tradeoffs between speed and accuracy is often used. The variation in decision-making strategy used by experienced decision makers results in selection of the first option that satisfies immediate requirements as compared to more analytical or procedural decision making approaches (Klein, 1993; Orasanu, 1999). The NDM approach calls for this phenomenon to be studied within meaningful contexts using participants with a range of experience in the domain studied as distinguished from normative models of decision making that originated for use in economics and statistics.

The naturalistic, or adaptive, decision-making strategy is a more useful approach to studying decision making under stress in aviation than traditional research offered with normative analytical strategies (e.g., Baysian probability theory). NDM provides a framework for meaningful, relevant research structure and interpretation of study results. The following criteria have been established for using the NDM approach for research: (a) context rich circumstances, (b) the use of expert participants, (c) the purpose of research is to discover strategies rather than trying to detect deviations from rational

standard, and (d) the locus of interest within decision episode includes situational awareness rather than a restriction to the choice alone (Zsambock 1993, 1997).

A number of NDM theories have been examined in military and aircrew settings. Decision making analyses (Cohen & Freeman, 1996; Lipshitz 1995; Sarter & Woods, 1997; Kuperman, 1998; Leedom, Adelman, Murphy, 1998) have been conducted to gain insight into how military decision makers reacted to situations and generated responses, and to determine what kind of information was seen as significant to the decisions. Several models of naturalistic decision making strategy have been advanced in the last decade that assist in explaining real-world decision-making. Naturalistic decision making models and theories maintain that the decision maker's expertise plays a central role in recognizing that a problem exists, in shaping the problem and in responding to the problem (Zsambok, 1997).

Naturalistic Decision Making (NDM) Models

Naturalistic decision models are used to investigate, describe, and predict decision-making styles across domains and provide the theoretical framework for this study. Of primary interest is work by Klein (1989,1993,1997) that acknowledges use of other decision styles, such as analytic or option comparison, but emphasizes that these are less effective in time critical or high stress situations. As Klein points out, the issue is "how people develop and use experience, and the types of strategies that are adapted to take advantage of experience" (2000, p.165). Klein (1997) found that people tend to follow what appears to be their "instincts" or their "gut reaction" in time-critical, unfamiliar situations in order to save time and effort in situational assessment and in comparing options when time is short and stakes are high.

Recognition-Primed Decision (RPD) model. The dominant model used to examine naturalistic decision making is the Recognition-Primed (RPD) model (Klein et al., 1986) that underscores the importance of domain-specific knowledge or experience required to generate and evaluate an effective course of action in a time-critical situation. This three step descriptive model of naturalistic decision making focuses on situational assessment required to accurately classify an unfamiliar situation rather than trying to determine which option will most likely result in a successful outcome. In the first step, once the problem has been appropriately classified the problem solver can then look for patterns or similarities to other situations in which he or she was able to negotiate a successful outcome.

The RPD model traces the use of the decision maker's experience in guiding this initial step and focuses on the decision maker's expertise of domain knowledge (i.e., critical cues and causal factors). This model differs from the approach taken in the classical model of decision making which requires the decision maker to "decompose the situation into basic elements and perform analyses and calculations on the elements. The model departs most sharply from the majority of classical models of decision making in its attempt to trace the use of experience" (Beach, Chi, Klein, Smith & Vicente, 1997, p. 30).

The adaptation by the decision-maker to a naturalistic approach serves to reduce information overload, confusion, and assists in establishing accurate expectations. The second step of the RPD model is to select from a serial presentation a solution that will work. The final step in the model requires a mental rehearsal of the action to identify potential problems (Lipshitz, 1995). In this model, rule-based decisions require the least

amount of cognitive work while multiple responses that must be evaluated in light of constraints and outcomes require more work (Payne, Bettman & Johnson, 1993). The greatest amounts of cognitive work are required if no response options are available and a response must be invented and evaluated for accuracy. This RPD model provides a frame for identifying weak links and biases in aviation decision-making processes.

Recognitional/Metacognitive (R/M) model. Another area of analysis in decision making is concerned with the way people characterize circumstances or assess situations. Cohen, Freeman, & Wolf (1996) argue that the standard normative approaches of assessing outcomes on the dimensions of its subjected value or utility and its perceived likelihood of occurrence or subjective probability, as represented by the subjective utility theory or Baysian probability theory, miss the capability of decision makers to construct stories to use their experience to consider pieces of evidence in context. Rather, the assessment of information is weighed in the context of the plausibility of a self-constructed "story" rather than taking the average of the weighted items to construct probability estimates that are problematical and time consuming to use. The Recognitional/Metacognitive (R/M) model serves as an extension of the descriptive RPD model by adding another framework in which to study naturalistic decision making, "metarecognition is a cluster of skills that support and go beyond the recognitional processes in situational assessment..." (Cohen, Freeman & Thompson, 1997, p. 258). The R/M model integrates meta-level controlled recognitional schemas used by decision makers to self critique their mental processes (i.e., assess response for problems of incompleteness, unreliability, or conflict), correct, and apply a "quick test" to give

insight into adjustments they may need to make to take advantage of available resources and opportunities (Cohen & Freeman, 1996; Cohen et al., 1997).

Aviation Decision Process (ADP) model. Orasanu and Fischer (1997) developed an Aviation Decision Process (ADP) model, is also based on RPD model, which served as a frame for analyzing crew performance in the simulator. The ADP model is a conceptual framework that allows "for the prediction of which decisions demand the greatest amount of cognitive work, and where decision errors are most likely" (Orasanu, 1997, p.49). This model consists of two major components: ambiguity in situational assessment and uncertainty in choosing a course of action.

Using a combination of observations of flight crews performing in high fidelity simulators and the Aviation Safety Reporting System database, Orasanu and Fischer (1997) conducted research on the relationship between stressors, decision errors, and accidents. The investigation integrated decision event data from major commercial airline mishap analyses from 1982 through 1997 with decision strategy in context data from two separate simulator studies. The result of this effort was the creation of a decision event taxonomy that includes only components that are observable in crew performance. Orasanu (1997) concludes that the combined study analysis discovered that more effective crews were more flexible in their "application of a varied repertoire of strategies" while less effective crews "did not appear to distinguish among the various types of decisions, applying the same strategies in all cases regardless of variations in their demands" (p. 356). The Aviation Decision model is tailored to this taxonomy and includes a range of decision-making strategies from "rapid intuitive decisions to

analytic, option comparison, and creation of novel solutions for unfamiliar problems" (Flin, p. 2, 1998).

Experience and Performance Differences

Research has indicated a need to identify the effect of different experience levels on team behavior and performance because studies done to extract training needs from individual members has been unreliable depending on the experience level a team member has with a particular behavior. "Research grounded in normative decision models tends to ignore the enormous power conferred by domain expertise" (Orasanu, 1993, p. 152) that is critical in crew time and workload reduction. This finding is supported by Cohen & Freeman (1996) who discovered related evidence that experience in efficient encoding and retrieval of domain-specific information is critical to interpretation of vague/nondiagnostic cues along with cues that may also be conflicting or difficult to interpret in context. Additionally, research has shown that metacognitive skills of experts are superior to novices. These skills include self-monitoring skills, better structure of domain knowledge (Patel & Groen, 1991) as well as the superior ability to perceive meaningful patterns and to retrieve domain-relevant facts (Glaser & Chi, 1988). Experts are more likely to terminate a problem solving strategy that is not meeting situational requirements (Larkin, 1983) because they are using their highly developed metacognitive skills.

Comparing Expert and Novice Performance

Orasanu and Fischer (1992) found that more effective commercial airline crews reduced communications as their workload demand increased and when commands were issued the content of the communication was directed towards future events. In another

study on the effect of experience level on decision outcomes of commercial airline pilots Stokes, Kemper and Kite (1997) used desktop flight simulation and cognitive testing to test the adequacy of traditional views of decision making in aviation. This research found that although experienced commercial pilots generated 30% more action alternatives than novice pilots and were far more likely (71% of the time) to carry out the first option that satisfied the circumstance, the experience level of the pilot was not the best predictor of decision-making performance in the scenario; rather it was the number of relevant cues detected.

Prince, Hartel, and Salas (1992) conducted an empirical study using military aviators. Thirty crews of undergraduate Naval aviators, of different experience levels, were used to investigate the relationship between team decision making-strategies and performance. There were two major findings of this investigation that are most relevant to this study. First, some crews tended to use the same strategy for all decisions they had to make in the cockpit; usually these were the least experienced crews. Second, the same strategy for decision-making is not equally effective for all decisions.

Team/Aircrew Performance and Evaluation

Although decision-making data is the primary focus of this study, it cannot be researched in isolation. Most successful decision-making strategies in organizations are dependent on assessing the current situation and determining/predicting which information will prove most useful in meeting immediate goals (Janis, 1989). Similarly, the interplay between decision-making, situational assessment, and crew coordination are primary dynamics in safety of flight, management of emergencies and critical decisions in the cockpit. In aviation contexts it is essential to consider the process of team

coordination as it affects cockpit situational awareness, shared problem models, decision-making processes and decision outcomes. Effective crew communication and coordination are integral to the interpretation, exchange, and use of situational cues and are a major focus for the collection and analysis of the data related to flight crew performance.

Results of a meta-analysis (Hartel, Smith & Prince, 1991) of Navy and Marine Corps mishaps that occurred between 1980-1990 revealed that decision-making errors contributed to 188 mishaps, problems in situational awareness contributed to 229 mishaps, and crew coordination errors contributed to 316 mishaps during the 10 year period covered by the meta analysis are summarized in Table 1.

Table 1.

Meta Analysis of Class "A" Naval Aviation Mishaps (1980-1990)

Type of Error	**Crew Coordination**	**Situational Awareness**	**Decision Making**
Contributing Factor in Percentage of Total Class "A" Mishaps	80%	60%	50%

Additionally, research conducted by the Naval Air Warfare Center examined these constructs and determined that (a) training in these complex cognitive, advanced team skills needed to be mission and context specific and that (b) these skills were perishable, thus requiring refresher training. Cannon-Bowers, Tannenbaum, Salas and Volpe (1993) conclude that theorizing about team performance and training has not yielded generalizable principles that have practical application. This is so partly because of the complexity of the team arena and in part because of the number of variables and

constructs that must be considered in the study of teams in their dynamic work environments.

Approaches to Team Research

Several empirical approaches to increasing knowledge of team behavior and strategies in measuring team performance are briefly discussed below as they relate to the theme of understanding and measuring team performance. The first two studies examined the use cognitive approaches to investigate links between team members' thinking processes and perceptions of team behavior importance. The third and fourth studies addressed the need for further theoretical development through applied research with a focus on the construct validity of measures of teamwork. This research uses military aircrews in simulated missions to address requirements to clarify and measure those process variables that are important to team functioning. A similar series of studies conducted with United States Air Force aircrews focused on measuring a wide range of individual and team processes as they relate to mission outcomes. Other research methodology and data analyses techniques discussed were employed in investigating team performance in incident management and modeling experiments that generated and validated mathematical models to predict optimal fusion of information for team performance are presented.

Team member experience levels and teamwork knowledge. A study by Rentsch, Heffner & Duffy (1994) explored the relationship between teamwork knowledge, or teamwork schemas, and team experience. This research investigated team members with varying levels of experience to determine whether this variable of experience played a role in teamwork conceptualization and subsequent team behavior. The methodologies

used in this investigation of teamwork knowledge in relation to teamwork experience were nontraditional to the organizational science literature. Although the methods used were "based on methods accepted by cognitive and education scientists who study schemas" (p. 455) practical implications from this study, which are "consistent with contemporary expert-novice literature" (Rentsch et al., 1994, p. 450), suggests that team training should be designed and delivered based on a person's prior teamwork knowledge and experiences. Additionally, the authors conclude that organizational managers and team leaders need to consider potential team members' experience levels when making team assignments. These authors deduced that if team members think about work differently depending on their level of experience, then team training should be focused toward team experience level. One of the limitations of this study that has implications for future research is that a median split was used to differentiate higher and lower self-reported team experience levels of participants. Researchers may need to identify other means to measure experience levels of individuals other than assignment to, and time spent on, various teams.

The team literature reviewed recognizes an established need to institute metrics for identifying critical team training needs. The purpose of the Baker and Salas (1996) study was to investigate the effects of experience on perceptions of team behavior importance in accomplishing a mission. In their literature review Baker and Salas found that "the effects of experience have been documented for individual tasks . . . but research has yet to investigate the extent to which team member experience affects perceptions of team behavior importance" (Baker & Salas, 1996, p. 238). In the Baker and Salas study with aviators, five dimensions of team behavior commonly used in task

and team performance analysis (i.e., criticality of error, difficulty, time spent, difficulty in learning, and importance for training) were rated by pilots on questionnaires developed to reflect specific teamwork requirements. These requirements were previously generated from a series of critical incident interviews with military pilots. Respondents included U.S. Navy instructor and student pilots randomly selected from a primary flight training organization. Pilots with varying degrees of flight experience used a seven-point relative rating scale to rate the five dimensions as they related to team behavior in their aircraft. Data were analyzed by correlating each behavioral dimension (dependent variables) with the pilot's overall importance rating. Job experience served as the independent variable.

In general, results of the study support the hypothesis that experience level would factor into team members' emphasis in the importance of various team dimensions. Less experienced crewmembers placed greater emphasis on difficulty on performing a team behavior while more experienced crewmembers judged time spent performing a behavior to have more criticality in overall team behavior importance. However, it is important to note that the less experienced aviators emphasized difficulty in performing a task rather than learning the task as posited by the first hypothesis in this research. These results suggest, "different team behaviors should be the target of training, depending on the experience level of the trainees" (Baker & Salas, 1996, p. 243).

Team performance measurement. A study by Reinartz (1993) explored issues in the development of methodology and analysis techniques to study how teams cope with multi-fault incidents. To achieve the aim of the team behavior research Reinartz explained that "it was necessary to develop both (a) a methodology to obtain behavioral

data and (b) analysis techniques to elucidate knowledge about this behavior which could be applied to the design of control rooms, procedures, and training systems" (1993, p. 1). In preparing the experimental design, a literature review included identification of the methodological approaches to studying complex decision making, procedural tasks, and protocols for data collection in a nuclear power plant. Actual workplace observations were used to create complex procedural scenarios for use in the experiment. During the experiment, scenarios were videotaped to ensure all data were collected. Additionally, to ensure a robust cognitive process trail, operators verbalized procedures that might have interfered with task-related interaction and led to confusion among team members. The verbal protocols collected from the operators as they performed their work greatly assisted in the analysis of teams' cognitive processes (i.e., planning, decision-making and problem solving) related to coping behavior. However, Reinartz discovered that sub-problem areas (e.g., defining team strategies, information choices and obstacles) were difficult to define with the data analysis approaches used. So through what the author describes as "... an iterative process of becoming more familiar with data, especially technical aspects of an incident handling" (1993, p. 7) a clearer direction to analysis evolved. Determining specific team behaviors of interest and their levels of description were accomplished using an "heuristic strategy of an ad hoc level of description and making use of all the discernible and recognizably useful information" (p. 7). A hierarchical task analysis was then developed and used to build up hierarchical team goal structures. This methodology, in concert with videotaping scenarios, assisted in examining individual and team activities in a chronological order. Reinartz also discovered that recording start and end times of operator behavior that added a

significant cognitive contribution to the team's subgoal/goal revealed differences in operators' strategies in handling a complex problem. This unconventional approach was found more useful than previous methodologies in documenting team behavior in complex non-procedural situations.

The same concern for the precision required in understanding the nature of effective team performance was addressed in a series of experiments designed and carried out by Kleinman, Luh, Pattipati, and Serfaty (1992). A major objective of this endeavor was to show that the normative-descriptive modeling approach to decision making, previously used in static situations, is a valid and powerful method for building quantitative models for team performance.

The first of a series of experiments was designed using two independent variables that the normative model predicted would affect the information-combining processing. U.S. Navy shipboard tactical teams were used to record activities that included the blending and integration of different internal and external sources of information. Data analysis by Kleinman et al. (1992) identified four cognitive biases in the team information fusion process that are useful in accurately predicting how team members weigh and combine sequential information from distributed sources of varying quality. Using mathematical models team member biases were quantified and integrated into a normative-descriptive model that predicted the actual experimental data far better than the normative model did. As the researchers concluded, there was a significant amount of knowledge gained in the insight into the human decision-making process by altering the normative model to capture human behavior. The subject's information sources,

including other team members, were received by the subjects with varying degrees of bias.

The findings of this research indicate a need to make team members aware of their potential to use poor decision-making strategies through faulty information integration. As reported by Kleinman, et al. (1992), subjects in groups of two and three in these studies had a pattern of: (a) consistently overweighing the most recent information; (b) continually placing some weight on prior knowledge; (c) not discounting common prior knowledge in communication updates; and (d) undervaluing the information received from their partner(s). These findings are important in providing a framework for further research in studying limitations and biases that may limit decision-makers.

Although data collection on team performance is usually restricted to simulated environments and controlled settings, this has not been as much of a restraint to research in team performance as the lack of measurements available to validate applicable team-related theories and models and decision strategies. According to Fowlkes, Lane, Salas, Franz and Oser (1994) the "slow headway made in understanding team performance may be attributed in large part to the lack of sound measurement approaches. Significant advancement requires theoretically based and psychometrically sound methods of observing and quantifying team performance" (p. 48).

The purpose of the team process measurement experiment by Brannick, Prince, Prince, and Salas (1995) was to document an approach used to develop and evaluate a set of tools to measure team process. Process was chosen over outcomes because, "in theory, process variables can be altered or effectively managed more easily than outcome

variables" (p. 642). In effect, process variables may present a "truer, richer picture of team functioning than outcome variables and they show promise for improving team functioning" (p. 641). The three central questions of this study were: (a) can judges provide psychometrically sound evaluations of teamwork (b) can those evaluations be correlated with more traditional expert evaluation, and (c) do such evaluations show good convergent and discriminant validity over occasions (i.e., scenarios)?

Participants in this study were 51 teams of Navy instructor and student pilots who were assigned randomly to one of 18 graduate student judges who evaluated crew coordination in each team in two simulated flight scenarios. The process indices used in this research were communication, cooperation (cohesion) and coordination as they are related to team effectiveness. In addition to subject matter expert scores two other methods of data collection were used. Two forms were also created and used in an attempt to attain a high degree of convergent and discriminant validity across judges. These forms were used to assign observable behaviors to specific dimensions and link behaviors in a dimension to evaluations of the behavior.

Data analysis was conducted to compare instructor and student pilot performance in the two scenarios using individual two-sample independent *t*-tests. Results of aircrew scores showed the means of instructor pilot ratings were generally larger than the mean ratings given to the student pilots in the same scenario and between scenarios. Although researchers refer to the rating difference as "significant", they did not report the magnitude of the statistic of difference between the two sample means that would justify rejecting the null hypothesis.

Data-based training. Evaluation of aircrews in a simulator by experts using evaluative criteria has been validated in many aviation-related studies. Most recently, this methodology has assisted in finding major effects between aircrew coordination and performance processes and mission outcomes in several aircrew studies in the United States Air Force (Spiker, et al, 1996; Nullmeyer et al., 1999).

Several U. S. Air Force-sponsored studies using a behavior-based and data driven approach have begun to pay major dividends for tactical military aircrew training. An Air Force sponsored experimental study conducted by Thompson, Tourville, Spiker, and Nullmeyer (1999) observed mission qualified crews as they planned and executed a mission in a high fidelity simulator. The purpose of this study was to examine the relationship between Cockpit Resource Management (CRM) skills (which include situational awareness and decision-making) and successful mission performance. Findings of this study using sixteen experienced MH-53J special operations aircrews were consistent with previous U.S. Air Force studies in various platforms such as the C-5 (Spiker, Tourville, Bragger, Dowdy, & Nullmeyer; 1999) and B-52 (Thorton, Kaempf, Zeller, & McAnulty, 1992) that represented an advancement in a practical descriptive approach to investigating and documenting aircrew behavior and performance. In the Thompson et al. study (1999), sixteen fleet crews' CRM behaviors and mission performance were independently rated by two Subject Matter Experts (SMEs) during specific mission phases using behaviorally anchored rating scales (BARS). Overall CRM ratings were then correlated with overall mission performance ratings to assess the role of CRM in mission performance. This study adds to the previous work that identified relevant crew CRM behaviors as different for diverse aircraft and mission

performance. Although the research methodology was weakened by lack of inter-rater reliability since only one rater was used for each criterion within the CRM and performance variables, the data collection instruments used are a valuable addition to both training and research methodology. Evaluation instruments defined specific aircrew behaviors at different phases of flight/the mission that consistently and reliably predict effective flight and mission accomplishment. The findings in this study were consistent with previous research using military aviators as study subjects. The conclusions of the study that have training and operational importance are: (a) CRM can be measured and analyzed when defined in terms of measurable behaviors, (b) CRM and mission performance are highly related (75% of variance accounted for and statistically reliable ($p < .001$), (c) the quality of the mission preparation predicts performance during mission execution, and (d) the specific behaviors associated with high or low CRM ratings were typically not covered in traditional Air Force training programs.

Summary

As discussed, the literature review describes a wide range of perspectives on designing and conducting decision making research in a team environment. The combination of theoretical frameworks and methodologies for studying the effect of experience on performance has been advanced but has only begun to meet the challenges of studying individuals and teams in their dynamic work environments. It is apparent that data collection methodologies and analytical techniques need to be improved and new ones generated to meet further research requirements as well as to generate effective training in adaptive decision making.

CHAPTER THREE

METHODOLOGY

Introduction

The purpose of this chapter is to describe in more detail than Chapter 1 the data collection and analytical strategies for this mixed design comparative case study. The study design participants, procedures, instruments, and data analyses are discussed. Techniques of data management are described, as are statistical assumptions and analytical techniques.

Design and Rationale

The design of this study employed both quantitative and qualitative methodologies to gain insight into themes that emerged in aircrew/team decision-making strategies and performance. This comparative case study approach integrated data collection and analysis strategies. Theories of naturalistic and aviation decision making processes, expert/novice performance, and team training literature guided the research.

Quantitative data analysis demonstrated the differences in novice and experienced aircrew process and outcome performance that were then decomposed descriptively. This strategy provided for a richer understanding of the data and is used extensively by behavioral and social scientists (Tashakkori & Teddlie, 1998). The focus of data analysis was on the effects of aircrew prior knowledge and experience in the detection, assessment, and risk management of atypical situations in-varying degrees of uncertainty and time constraints.

This study used a within case analysis, as typical for multiple case treatment, and provided a thematic description of each crew group experience. As expected, each successive case analysis informed a deeper understanding of subsequent analyses so an iterative approach was undertaken that guided a thematic cross-case analysis. The qualitative structure of this study centered on the use of an inquiry method that was distinctively suited to reveal the nature of the circumstances and the thoughts experienced by aircrews. The context rich environment forced aircrews to assess risk and manage non-standard approaches to problem resolution by recognizing and dealing with ambiguous cues and multiple failures in a series of dynamic situations.

Data collection and analysis triangulation used a diversity of sampling and analytical strategies by multiple observers, interviewers, and analysts. Aircrew experience level served as the dependent variable for the study. The independent variables were process ratings in seven observable scenario tasks as well as outcome performance rankings based on final disposition of the crew and the aircraft.

The following study hypotheses and research questions framed the relationship between aircrew experience and phenomena related to effective aircrew situational assessment and decision making under varying degrees of uncertainty and time constraints:

Hypothesis 1: Experienced aircrews will receive better process ratings (i.e., technical skills, workflow, information sharing, consensus building, operational risk assessment and management and back-up routine) than novices in handling a specific in-flight emergency situation involving uncertainty.

Hypothesis 2: Decisions (i.e., outcome rankings) made by experienced aircrews in an uncertain situation involving a specific in-flight emergency are rated as being of higher quality than those made by novice aircrews.

Hypothesis 3: Process ratings are correlated with the decision outcome ratings for a specific in-flight emergency involving uncertainty.

Research Question 1: What adaptive recognitional/metacognitive decision-making strategies emerge from aircrews in a specific in-flight emergency involving uncertainty?

Research Question 2: In what way do adaptive/metacognitive decision-making patterns differ among successful and less than successful aircrews challenged with a specific in-flight emergency involving uncertainty?

Research Question 3: What characteristics/factors seem to define the most successful aircrew outcomes in a specific in-flight emergency involving uncertainty?

Participants

This research study was conducted in a field setting at a naval aviation training facility using active duty aircrews in a high fidelity flight simulator. The target populations for this study were experienced and novice aviators (pilots and naval flight officers (NFO's) who serve as copilot tactical coordinators (COTAC's)). This study was limited to an S-3B aircraft non-tactical mission using only two of the three crewmembers. (The third crewmember serves as a weapons system officer (TACCO) in tactical environments.) For the purpose of this study, the non-essential TACCO position was unoccupied. Since the study scenario did not involve a tactical mission element the primary responsibility for the Naval Flight Officer, or COTAC, was to serve as the navigator and to provide safety of flight back up for the pilot.

Sixteen 2-person crews (1 pilot and 1 COTAC) were represented in the final data set in this study. Eight crews were made up of students nearing the end of initial S-3B qualification training. The other 8 crews were experienced aviators undergoing annual refresher training. Operational flight trainer and student availability were the prime considerations for novice crew participation. Trainer availability and annual flight standardization (i.e., NATOPS check) requirements were the main considerations of crews paired and scheduled from fleet squadrons.

Novice pairings of aviators for the non-syllabus event were in the same class group and were of equal military rank with similar flight hour levels. Experienced participants in the study represented four fleet squadrons and one aircraft carrier ship's company staff billet. Fleet squadrons were requested to pair the most experienced pilots and COTAC's. Most experienced participants used the study scenario as part of their annual NATOPS re-qualification training requirements. Fleet pairings of aviators generally had higher ranking, more experienced pilots paired with lower ranking COTAC's and all but one crew of fleet experienced participants were crewed with members of their own squadron.

Attrition from the Study

All participants that completed the scenario agreed to remain in the study. However, two fleet crews that completed the study scenario and the debrief/interview were replaced in the data set because instructors allowed them to land on alternate runways at the home field. This option deviated from the scenario structure and intent to place the crews in an extremis naturalistic decision environment for the final scenario

event. In order to control for unequal sample sizes two additional experienced crews were obtained.

Participant Demographics

Study participants were drawn from the S-3 aviator population and ranged in age from 24 to 41 years old. All aircrews consisted of one pilot and one COTAC who had completed basic naval aviation flight training between 1983 and 2001. Novices had less than 100 hours in the S-3B. Experienced crews had between 700 and 5,000 flight hours in the S-3A/B. The S-3A was a previous variant of the S-3B modified to include enhanced tactical capabilities in the early 1990's. Although many of the "backend" systems were upgraded, there was little done to change the cockpit, the aircraft handling characteristics, or the "wings and engines" -related performance capabilities of the aircraft. Accordingly, senior pilots and COTAC's with experience in both the A and B models were considered to simply have a greater number of hours in the S-3 aircraft.

Combined crew flight hours of crew assignments for the study scenario are displayed in Table 2. Crews 1 through 8 were novice crews and crews 9 through 16 were fleet experienced crews.

Table 2

Paired Pilot and COTAC Crews' Combined S-3 Flight Time

Assigned Crew Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Combined Crew Flight Hours	90	88	52	70	183	180	130	54	3050	4200	1570	1630	4600	1600	1650	5650

The participants' total aircraft model flight hours, simulator hours, and recency of simulator and flight times are reported in Table A1, Appendix A. Table 3, below, presents the range and mean of S-3 flight time for both novice and experienced groups.

Table 3

Individual (Pilot and COTAC) and Combined Crew S-3 Flight Hours

Crew Level	Pilot Range of Hours	Pilot Mean Hours	COTAC Range of Hours	COTAC Mean Hours	Combined Crew Range	Combined Total Crew Mean
Novice	12-80	54	24-100	53	52-183	52
Experienced	650-3700	1569	730-2100	1425	1600-4600	1497

Representativeness of the Sample

The representativeness of both experienced and novice aircrews in the sample population reflect the general range of flight time in the S-3 community. The training command sample of novice aviators included 16 of the 75 students that annually complete the familiarization phase of instruction at the Fleet Replacement Squadron (FRS). The sixteen study participants represent approximately 20 % of novice crews in the S-3 community. Four of the ten S-3 fleet squadrons were represented in the study sample and the eight crews of two represented approximately 10 % of the total front seat aircrews (pilots and COTAC's) assigned to fleet squadrons in the S-3 community.

Instructor/Subject Matter Expert Raters

Observers who were experienced flight and simulator instructors rated pilot and COTAC crews' process performance. Subject Matter Experts (SME's), with at least five years instructor experience and between 1100 and 2700 hours of flight time in the S-3A and/or S-3B were personally selected to participate in this study based on their reputations as exceptional, "non-threatening" instructors by both student and peer evaluations that are conducted routinely at the Fleet Replacement Squadron (FRS). A

total of eight process raters (seven civilian and one military) used scenario-specific rating instrument to evaluate participants. At least two pilot and two NFO (i.e., COTAC) raters were used for each crew. Depending on availability, between one and four trained senior evaluators participated in the actual simulator event and debrief/interview. Other process raters evaluated aircrews from audio, visual and simulator instrument panel data collected on the computer debriefing system that recorded the data during the event.

In order to control for experimenter-expectancy bias a group of at least four observers (several crews had five observers) rated process and an independent group of four outcome raters were used for the final outcome ranking to decrease learning of influence techniques, to randomize expectancies, and to increase the generality of results. A fixed outcome data set for each crew were presented to the second set of independent raters (active duty and reserve) that were not involved in the study and were not assigned to squadrons of the study participants. Demographics for process raters are contained in Table A2 of Appendix A.

Protection of Human Subjects

Administrative approval for conducting and publishing this study was sought from cognizant authorities of eight organizations: (a) the university from which the doctoral degree will be granted, (b) the commander that oversees United States Pacific Fleet Air Forces, (c) the administrative authority for S-3B fleet squadrons and training simulators, and (d) the five squadrons from which participants in the study were assigned. Appendix B contains the approval to conduct the study from the United States Navy by the direction of the Commander of the Naval Air Force Pacific Fleet as well as

the approval of from the S-3 community leadership at North Island Naval air Station. Appendix C contains the participant consent form.

All participants were told they could opt out of the formal study after the simulator event and still obtain an "off the record" debrief. All crews that participated in the scenario agreed to remain in the study. Two crews were disqualified because instructors during the scenario event used non-standard procedures. The researcher did not receive any negative feedback on the conduct of the study from study participants or their respective commands. The data collection environment provided a safe, realistic venue to present and provide aircrews with multiple opportunities to assess and respond to critical safety of flight problems both during and after the simulator session.

Scenario Design Elements

Since it is usually optimal to have comparable degrees of detail and precision in the scenario design as are experienced in real world flight (Cook & Campbell, 1979) the study scenario was designed to represent a combination of actual atypical and emergency events that have or could occur in the S-3B aircraft. Because "decision making is highly contingent on the demands of the task" (Payne, 1982, p. 382) the final study scenario event design reflects a range of dynamics associated with task demand to elicit different types of problem solving and judgment. A variety of cues and situations were developed for the study scenario to reflect elements and factors that are potential issues in the S-3 crews' operational decision making environment. The naturalistic environment as described by Orasanu and Connolly (1995) may include several or all of the following: (a) ill-structured problems, (b) uncertain dynamics, (c) changing, ill-defined or

competing goals, (d) crew action/feedback loops, (e) time stress, (f) high risk, multiple players, and (g) consideration of organizational goals and norms.

Using the naturalistic decision making elements as guidance, the specific contexts for scenario tasks were obtained from a combination of (a) recent S-3B mishap and training data, (b) an S-3B cognitive task analysis for decision skills curriculum development, and (c) the framework developed by Hammond, Hamm, Grassia & Pearson (1997) that describes the cognitive properties that distinguish intuitive and analytical approaches to decision making. As recommended by Hammond et al. (1997) intuition-inducing task conditions were used to set the final decision apart from previous tasks. The intuition-generating characteristics that set the final decision environment apart from previous events were as follows:

- a shorter response period allowed
- a larger number of cues required
- simultaneous cues given
- perceptual rather than objective measurement of cues (i.e., ill-structured)
- a low rather than high decomposition of the task required
- no organizing principle available from standard operating procedures (SOP's), previous training, or experience (i.e., unstructured)
- a low rather than high certainty that the task(s) could be accomplished safely and within established procedures

A synopsis of the specific naturalistic decision elements and appropriate procedural, tactical, and strategic crew responses (S. K. Hunt, personal correspondence, December 10, 2001) associated with each of the seven scenario trigger cues in the study scenario is contained in Appendix D. The elements associated with each of the seven process ratings were developed and reviewed by four senior instructors. The review took

into consideration the instructors' knowledge of both experienced and novice aircrews' time to acknowledge recognition of problem cues, time to troubleshoot aircraft indications/respond to cues/information, and manage and coordinate accumulated and projected tasks.

The prototype scenario design was tested with two novice and two experienced aircrews to validate data collection categories, criteria, and protocols and included determination of: (a) the strength of trigger stimuli in the scenario events, (b) requirements for performance evaluation criteria, (c) validation of most pertinent aircrew and aircraft parameter outcome criteria, (d) post-hoc aircrew interview/debrief protocol and questions, (e) reliability of digital playback file capture, and (f) the most robust multivariate statistic for data analysis.

Based on the initial evaluation of the scenario, modifications were made to the scenario brief and instructor protocol guidelines to eliminate aircrews' selection of alternative courses of action (i.e., land on a different runway). Additionally, adjustments made in the scenario event stressors forced more rapid rates of data processing and decision making by aircrews in each subsequent scenario event.

Process Variable Definitions

The scenario process variables were defined in a joint effort between the researcher and senior instructors using previous S-3 mishaps, training trends, and prior cognitive and skill task analysis results used for decision skills integration into the FRS syllabus. Events were selected that induced or "triggered" procedural, analytical, and/or an adaptive intuition-inducing cognitive state based on the task properties and time

available. Table 4 summarizes task elements related to each of the seven scenario events observed for process scores.

Table 4

Scenario Event Trigger Conditions, Problem Structure, and Associated Task Requirements

Event Sequence	Cue /Trigger	Structure of Problem	Crew Task
1	Normal flight instruments	N/A	Takeoff under normal conditions
2	Illumination of flashing red Master Warning Light and No 1 engine Starter Light	Master Caution Panel Lights in direct line of vision of Starter Light. Well-defined procedures for safety of flight.	Complete emergency procedure using Checklist as reference
3	Checklist	Normal structure interrupted by Tower	Complete checklist. Dump fuel.
4	Oil pressure gage indication drops significantly in No. 2 engine	Ill-structured significant safety of flight issue with only remaining engine.	Retard throttle to IDLE and monitor engine. Crew unable to follow procedure that requires shut down of engine to preclude fire and possible loss of associated systems.
5	Deteriorating No. 2 engine (Attempts to re-light No. 1 engine)	Ill-structured under extremis	Consider contingencies for single-engine approach and go-around profiles. Monitor No. 2 engine. Selectively choose/abbreviate sections of checklists to comply in time available. Dump fuel if not done.
6	No arresting gear on only available runway	Ill-structured under extremis	Contingency planning- Brief loss of No. 2 engine. Brief wave-off and hook skip contingencies.
7	Fouled deck at intersection of runway	Novel and unstructured with immediate action required.	Crew must determine if they can land and stop by intersection.

Outcome Variable Definitions

Instructors that assisted with the scenario design defined outcome variables a priori. The variables that made up the final aircraft position and configuration "snap shot" data were used to identify priorities in the quality of the crews' execution of the final event decision. The simulator data sets obtained from the computer debriefing system display and data analysis software, reported in Appendix E, were used for the outcome ranking of all crews. The outcome data sets included: (a) airspeed, (b) altitude, (c) angle of attack, (d) flap position, (e) gear position, (f) hook position, (g) fuel remaining, (h) speed brake position, (i) emergency hydraulics pump (EHP), (j) visual speed indicator (VSI), (k) heading, and (l) geo position.

Final Scenario Design

The final scenario design established a simple baseline takeoff and departure procedural decision that could be compared to less structured situations requiring more complex decision making to resolve unclear and/or unstructured problems as the flight scenario progressed. The last of seven scenario events established a novel situation that required a creative problem-solving decision strategy. The seven process scores are based on the crews' recognition of and reactions to the events as sequenced in the study scenario. As described earlier, the task condition for the final event was distinguished by the greatest number of intuition-generating characteristics (i.e., lack of structure). The novelty of the truck in the runway on final approach leaves the crew with no standard procedures or rules to guide them in their response. The study scenario events are depicted on the scenario timeline in Figure 1.

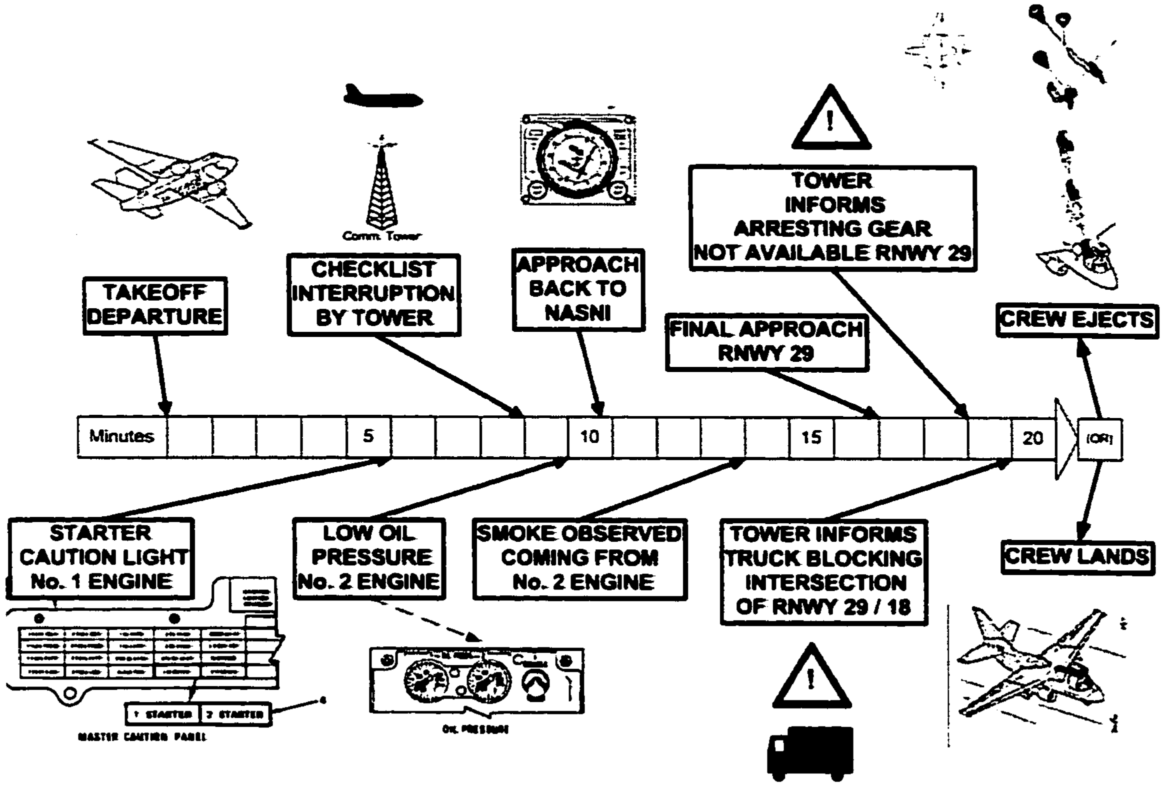

Figure 1. Scenario timeline and event flow.

Sequence of Procedures

Data were collected over a two-week period for novice crews and five-month period for fleet crews. Upon reporting for the scheduled simulator event, participants received a letter from the researcher with a brief explanation as to the nature of the study and the selection criteria for participants (i.e., experience level and no previous exposure to the study scenario). After discussing the consent form with the participants and upon receiving signed permission to use participant simulator and debrief data for study purposes, a personal background data form was given to the participants to complete.

Flight Simulator Scenario Protocol

Participants had 10 minutes to conduct a crewmember brief after the instructor gave them the conditions for the scenario. The 20-minute scenario event was conducted as outlined on the evaluation sheet. The study scenario was given under as similar as possible conditions in the flight simulator for all crews. The instructor/raters used the scenario event timeline as guidance for event cue initiation and for tower communications. Instructors used their discretion to respond to crew inquiries and to manage the seven scenario events within the 20-minute time allocation.

Data Collection and Replay Device

A prototype data collection device, customized for the S-3B operational flight trainer, digitally captured flight simulator data and is depicted in Figure 2. This system was used to digitally recall and display data from selected cockpit instruments and controls, a three-dimensional model of the aircraft, and an audio/video file of the aircrew.

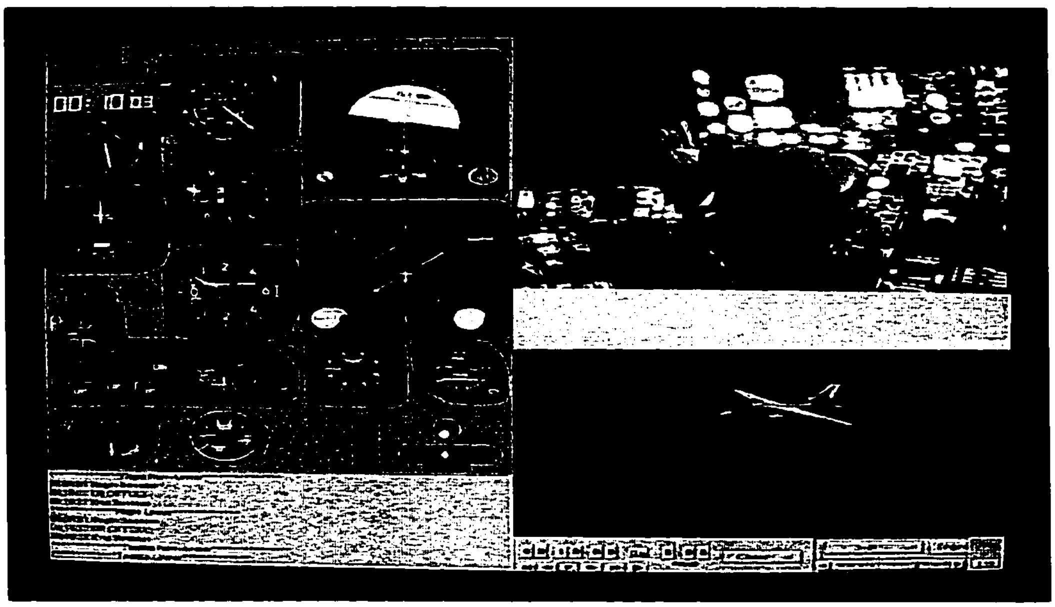

Figure 2. Digital debriefing system screen display.

Although the entire event was captured on a digital file, instructors were able to "mark" specific portions of the flight event for immediate access during the debrief to aid both instructors and participants in the recall of events and to support instructional points. Field notes recorded aircraft geographical position flight path patterns and served as a back-up to the digitally recorded files. After completion of the scenario the participants were escorted back to the briefing room and reminded that the debrief/interview would be audiotaped. The 40 to 60-minute tape recorded debrief/interview with the researcher and between one and up to four instructors, and up to three instructors, was conducted in a debrief room with a debriefing station to replay digital debrief files.

Debrief/Interview Protocol

During the debrief/interview, semi-structured questions were used to guide the debrief/interview. Crew and crewmember self-assessments were prompted using probe questions originated (Klein et al., 1986) to elicit specific information about situational awareness and decision elements. These questions were augmented by task-specific instructor and researcher queries and comments. Participants' post scenario interview/debrief data were recorded and transcribed by the researcher and coded a posteriori from interpretations of those data. Content analysis of decision processes revealed by the participants in the post-scenario interview/debrief focused on participant verbalization and self-assessment of their reconstructed experiences (both apparent and tacit). Baseline interview questions related to situational awareness issues and decision strategies are found in the Instructor Guide (Appendix F). There was no mention of process ratings during the interview/debrief. Field notes were taken during the debrief

by the researcher as a back up for the tape recording and to establish the position of the crewmember for transcription as their comments were audio taped.

During the debrief/interview relevant portions of the flight event "marked" by instructors was used to replay flight information (i.e., flight instrumentation, aircraft aspect) as well as communications and crew behaviors. The digital files enabled instructors and aircrews to access marked portions of the flight for recollection and/or clarification of event experiences. The recall of self-reported situational assessments and judgments were obtained on a tape recorder for both individual and crew interactions. At the completion of the debrief/interview, the tape recorder was turned off so comments by the researcher, instructors, and/or the participants could be made "off the record." The debrief concluded with delivery of copies of the signed consent forms to participants.

Procedures for Raters

Crew process performance data were obtained from operationalized study variables in the flight scenario. Raters used the criteria established for the process variables in the scenario instrument to guide their evaluations. Process raters also independently recorded general observations of the actual simulated event. Raters not present for the event watched the digital audio/visual files of the scenario session and recorded their process ratings and general comments on the evaluation sheet for each crew.

At the conclusion of process data collection for all crews a separate set of data related solely to aircraft parameters, configuration, and spatial orientation in the final "snapshot" of the event was created. The final outcome data sets for the 12 crews were de-identified of crew numbers and replaced with randomly assigned alphabetical

designators to eliminate a possible source of rater bias. These outcome data sets were then presented to an independent group of four outcome raters to rank order using their expert judgment. The final outcome ranking process took 90 minutes and was facilitated and observed by the researcher and one instructor. Each of the four subject matter expert raters first independently ranked crews on outcome data provided using a ranking sheet that required written justification for the order of their rankings. After individual rankings were completed a discussion of rating priorities and a multi-voting process was used to gain rater consensus for a collective judgment of the final outcome ranking of all crews. Outcome rater demographics and their independent and consensus outcome rankings of the sixteen aircrews are found in Appendix G.

Process Rating Instrument

Instructors used a scenario-specific evaluation instrument with a five-point ordinal rating scale to measure process variables. Independent crew process performance rating by expert observers was guided by criteria based on best practices associated with crew coordination and diffusion of tasks in events requiring sensitivity to a configuration of an entire profile of cues, safety of flight considerations, technical skills and decision making skills. The study rating instruments reflected the evaluation of group processes and outcomes for a multi-crew aircraft. Rating criteria were designed to ensure as much convergent and discriminant validity as possible across raters.

Process rating criteria were designed to generally capture multiple aspects of crew behavior and degrees of satisfactory workload sharing and operational risk management with consideration of safety of flight, adherence to standard operating procedures when practical, creative problem-solving, technical competency and crew

coordination in different types of problem solving contexts. Since group processes "are not simply a sum of individual processes (e.g., perception, attention, cognition) but are categorically different and include communication, information transfer, management processes, team problem solving and decision making" (Kanki, 1996, p. 136) team dimensions were incorporated into both the specific and common flight evaluation criteria.

The Team Dimensional Training (TDT) evaluation instrument validated with Navy aircrews (Smith-Jentsch, Johnston & Payne, 1998) was modified for use as the overall guidance for team rating criteria required for all study variables. The six cockpit resource management (CRM) elements used to measure U.S. Air Force C-5 aircrew proficiency (Spiker, et al., 1999) were also integrated into the overall aircrew process performance criteria as well. Additionally, context-specific best practice standards were identified in the rating instrument (Appendix F) to promote discriminant validity. These standards were included in the scenario instructor guide timeline to direct raters to established criteria for standard crew behaviors expected during each of the seven decision events.

Analysis of Reliability of Instructor Ratings

Snedecor's analysis of variance formula, derived from Fisher's work on intra-class correlation (Ebel, 1951) was used to estimate the reliability for instructors' averaged ratings of aircrew process performance, to validate rater inter-rater reliability, as well as to validate the measurement instrument designed specifically for the study scenario. Since individual ratings are less reliable than composite ratings the estimate of reliability of average ratings instead of individual ratings was used in this study to compute

statistical significance of data. The reliability of those averaged ratings was of primary concern. Given that compared averages came from different groups of four to five raters the between raters variance was included as part of the error terms. Appendix H contains individual scenario event and summed process scores for all crews.

Of the seven process ratings listed in Table 5, it can be seen that in five areas the averaged agreement coefficient of raters (N=8) shows a high degree of agreement within the groups of raters. Group averaged ratings are close to or exceed the .80 level of significance generally considered the standard for acceptable inter-rater reliability (Cronbach, 1990). The two variables, *take-off/departure* and *checklists interrupted*, do not meet Cronbach's criterion for acceptable inter-rater reliability. In the *take-off/departure* event there was virtually no variability in the ratings across either raters or crews. For the second variable, *checklists interrupted*, this scenario component represented a broad range of tasks interrelated with checklist discipline specific to this scenario segment.

Table 5

Inter-rater Reliability Correlations of Instructor Process Ratings of Aircrews

Process Evaluated	Take-off/ Departure	Abnornal Engine #1	Checklists Interrupted	Low Oil #2 Eng.	Approach Plan	De-rigged Gear Plan	Fouled Deck Plan
Raters' r	.14	.84	.61	.89	.76	.76	.83

Data Collection

Data collected for this study included: (a) participant and aircrew demographic information, (b) aircrew process performance ratings during the simulated flight event, (c) simulator data including aircraft configuration and temporal measures of aircrew recognition of low oil pressure indication(s), (d) post-scenario participant(s) self-reported performance issues and strategies from debrief/interview session, (e) simulator flight data and audio/video recordings of crew performance, and (f) researcher and subject matter expert field notes. These data were analyzed for any significant relationships between participant crew decision-making strategies and their final crew outcome ranks.

Cross-Case Comparison Analysis Strategy

Analysis of qualitative data began with a content analysis of the sixteen aircrews' debrief/interview transcript data. This content analysis was then merged with instructor observation comments, researcher field notes, raw and summed process ratings, and outcome performance rank-ordered data to conduct a secondary concurrent integrated analysis. This initial exploratory analysis included an "extreme case" comparative analysis of the overall highest and lowest performing aircrews in both experienced and novice groups and identified the emerging constructs or themes associated with individual and team attributes related to decision-making. As described by Tashakkori and Teddlie, this type of analysis strategy results in an initial "identification of groups of individuals who are similar to each other in some respect" (1998, p. 133). An additional strategy for interpreting these data included a comparison of participant aircrews' heuristic rules to "optimal" rules provided by subject matter experts since investigation of heuristic permits the examination of discrepancies between actual and optimal

behavior which then raises questions regarding why such discrepancies exist (Einhorn, 1980).

Statistical Assumptions

The goal of the statistical tests was to explore for possibly demonstratively different performance processes and results by comparing the sample means of data collected between and among the novice and experienced groups. The goal of the statistical tests was to validate the assumption that the crews with the greater number of flight hours would obtain better process and outcome ratings than the novice crews. The alpha level used was at a 95-percent confidence interval to determine if relationships existed among identified study variables (i.e., crew experience levels, crew process measurements, and effectiveness measurements of scenario outcome).

Chronology of Analysis

Inferences from the initial qualitative examination of data expanded with statistical analyses and further qualitative investigation. Data patterns and relationships were explored using several frames of reference. First, descriptive and inferential statistical analyses (MANOVA) explored the effects of crew experience on process ratings and outcome rankings. The relationship between process and outcome scores was also assessed. Second, comparative analyses were completed with the data related to the central propositions of the study: (a) a non-parametric and multivariate analysis (multiple ANOVA's) provided construct validation and confirmed and expanded the relational inferences made from the initial qualitative analysis, and (b) correlational analysis using the "Pearson product moment" explored data to identify process characteristics that seemed to be related to outcome. Third, further qualitative

decomposition of all crew data related to process and outcome was conducted to investigate supportive descriptive trends to determine if effects were limited to some subset of all areas considered in the study (e.g., Were scores uniformly pointing in the same direction or did some appear more important than others?). Fourth, selected case studies representing high, mid-range and low performing crews in both process and outcome ranks were compared for differences in the integrative complexity of each aircrews' decision-related metacognitive activity (i.e., information gathering/processing and decision making), communications, etc. Participant remarks during debrief interviews concerning their aviation skills (e.g., procedural, representational, flight management, decision making, etc.) related to their reconstructed reality of the scenario process elements were used to confirm and expand the inferences derived from previous data analyses.

Fifth, a more in-depth investigation of the crews receiving the highest, mid-range, and lowest process and outcome rankings was conducted to seek out differences in crew decision making characteristics and functioning. A crew transcript of major decision points was developed from audio/video files of each crew's flight performance and was then compared to the debrief/interview data transcript for a more comprehensive analysis of team behaviors in coordination and communication of strategic, tactical and procedural decision making. In conclusion, a variety of reporting techniques were used to establish the best combination of exhibits to portray the complexities of the participants reactions and interactions during the scenario, the interrelationships among variables, and the comparison of study findings to existing naturalistic decision making theories and selected aviation decision models.

Limitations of the Study

Overall, the process evaluation instrument reflects both real crew differences and measurement fluctuations, although there was little variability in the *takeoff/departure* element of the process scores. This was due to a lack of defined measurements to accurately assess variations in this baseline measurement procedure. As far as the reliability of the raters to accurately measure crew performance, the estimated aggregate internal consistency (mean reliability) for the group of raters on 6 of the 7 process ratings was at or very close to an acceptable level (.80).

Because error variances affect both the reliability and validity of measures, scoring methods for the instrument, characteristics of the participants (lack of preparation, anxiety), and/or lack of precision in the data collection instrument may have contributed to measurement error. Although outcome raters used their individual judgment for rank ordering crews before a consensus vote, there may have been pressure to conform their beliefs in line with those aviators who were more senior in the group.

One threat to internal validity in this study may have been the criteria established for the scoring of two of the seven items in scenario the measurement tool. As discussed previously, there was little variance in the *take-off /departure* item that served as a baseline for procedural compliance. Most crews on a normal take-off do not deviate from standard operating procedure. In another category, c*hecklist interrupt,* the construct was most probably too complex and expanded into more time in the phase of flight to be a single category because raters found more to comment on than was originally intended. Additionally, problems with the instrument may have occurred because instructors are unfamiliar with using criterion-based evaluation tools. Instructors typically use grade

sheets that simply list the skill/behavior to be observed without defined evaluation criteria. An ordinal scale of unsatisfactory, below average, average, and above average is used. Grading is generally non-standardized and moves from lenient to higher standards as the instructor gains an experience base in evaluating a particular event.

Summary

In recognition of the complexity of the issues to be studied a cross-case study comparison approach using both quantitative and qualitative methodologies was employed. Decision making process performance and execution measurement and analysis required a decomposed variable-oriented quantitative approach but also called for an investigation of individual cases to gain an understanding of decision process characteristics and related crew functioning.

A triangulation of data generation, collection sources, separation of process and evaluation criteria and raters, and analytical methodologies supported the reliability and validity of the study data and findings. More specifically, the design of this study included data generation and collection sources to include: (a) domain-specific tasks designed to elicit various decision strategies, (b) a criterion-based evaluation instrument to capture differences in crew process performance, (c) a simulator event process and outcome data collection and replay tool, (d) an interview protocol to generate crew self-refection and self-assessment on metacognitive and interactive processes, (e) observations of participants recorded in researcher and instructor field notes , and (f) judgment of outcome by independent raters.

Integration of the analysis methodologies supported the integrative evaluation of the data. Quantitative data was analyzed with descriptive and statistical tests to

investigate differences in crew process performance, execution of the final decision in the scenario (outcome), and the relationship between process performance and outcome ranking. Qualitative analysis incorporated cross-case meta analysis by flight hour and performance levels, an in-depth cross-case analysis of three process and three outcome levels of performance from various theoretical and practical perspectives.

CHAPTER 4

CROSS-CASE ANALYSIS

Introduction

Data related to each of the three hypotheses and three research questions are presented and discussed in five sections of this chapter in the following order: (a) aircrew process ratings related to Hypothesis 1; (b) aircrew outcome rankings related to Hypothesis 2; (c) the relationship of process and outcome ratings related to Hypothesis 3; (d) decision-making strategies related to the first two research questions, and (e) critical characteristics and factors that defined the most successful crew outcomes related to the third research question.

After a brief overview of the study simulator scenario, cross-case analyses of sixteen case studies are reported using expert ratings and rankings, instructor/ subject matter expert comments, instructor observations recorded during the simulator event, participant interview/debrief transcripts, and simulator flight data files with associated digital audio/video files of the sixteen aircrews. The quantitative analysis looked at how highly experienced crews differ from less experienced crews with regard to process performance and outcome rankings. The qualitative findings are presented collectively in a cross-case analysis between and among performance levels. The qualitative assessment focused on the differences in two main areas: the coordination of the decision making process between the crewmembers (including leadership, teamwork, information gathering and use), and the criteria essential to critical thinking processes and strategies associated with evaluating the situation and dealing with uncertainty (strategic focus, technical skills, risk assessment, etc.).

The differences in adaptation of crew behavior in decision-making strategies that emerged at various levels of process performance (high, mid-level, low rankings) are included along with common crew performance characteristics associated with high and low performing groups. Chapter 5 will interpret the crew data with respect to differing domain experience, cognitive effort, crew leadership, and crew coordination as considerations for selection of decision-making strategies and the execution of those decisions under conditions involving various degrees of uniqueness, uncertainty, and time limitations.

The Scenario

The demands of the process tasks are related to the task properties and the time constraints within the scenario. As the scenario progresses from a sequence of structured events to a succession of less structured events and then to a completely unstructured event, the appropriate response strategies change. The aircrews were exposed to seven process events in a good weather, nighttime scenario. The first event involved a normal takeoff and departure from the home field that served as a baseline procedural evolution. The second event presented a Starter Caution Light on the No. 1 engine that required a commonly practiced procedural response involving an intentional engine shutdown and an expeditious return to the field. The third event required the crew to prioritize checklists and apportion available time to complete high priority items while being interrupted by the control tower to be informed of another inbound emergency aircraft.

The fourth event began a series of more complex, unstructured situations as highlighted in Figure 3 below. Process event 4 began with the presentation of a low oil pressure indication on the only remaining engine. This cue was not as obvious as the previously introduced Starter Caution Light due to the ergonomics of the instrument panel (i.e., placement of the engine oil pressure gages, the relatively small size of the

engine oil pressure gages, and a conflicting warning light logic sequence which under certain circumstances does not visually or aurally draw attention to an oil system malfunction).

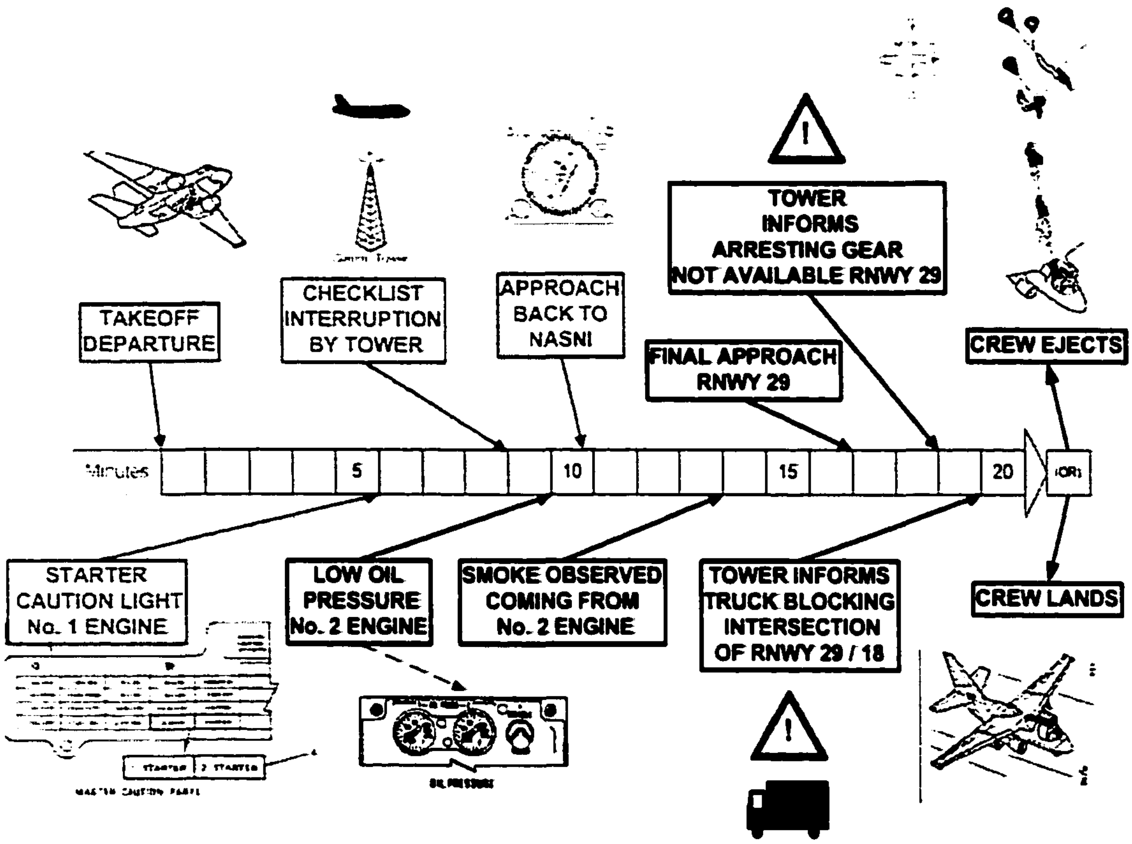

Figure 3. Highlighted unstructured scenario events starting with low oil pressure indication.

The fifth event involved preparations for a single-engine approach to the field. This required crewmembers to coordinate more checklists, dump fuel to reduce their landing weight, monitor the engine condition, work together to restart the previously secured No. 1 engine as a back-up to the deteriorating No. 2 engine, communicate with the control tower, plan for contingencies, and make related flight path adjustments under extremis conditions. During the sixth event, the control tower informs the crew that the

airfield arresting gear, required by standard operating procedures for a single-engine landing, is not available. The seventh and final event challenged the crew with an in-close "foul deck" call from the control tower while the crews were under an extremely high workload. This novel event presented crews with a set of circumstances that required them to execute a nearly instantaneous decision to either continue the approach or initiate a go-around (wave-off).

Quantitative Cross-Case Analysis Findings

Aircrew Process Ratings

Hypothesis 1 predicted that process ratings would be higher for experienced crews than novice crews. In general, experienced crews had higher process scores than novice crews. Figure 4 shows the difference between each crew's summed process score and the average process score across all 16 crews. Most crews followed the general pattern in the predicted direction of performance evaluation. Crews 1-8 are novice crews and crews 9-18 are fleet experienced crews.

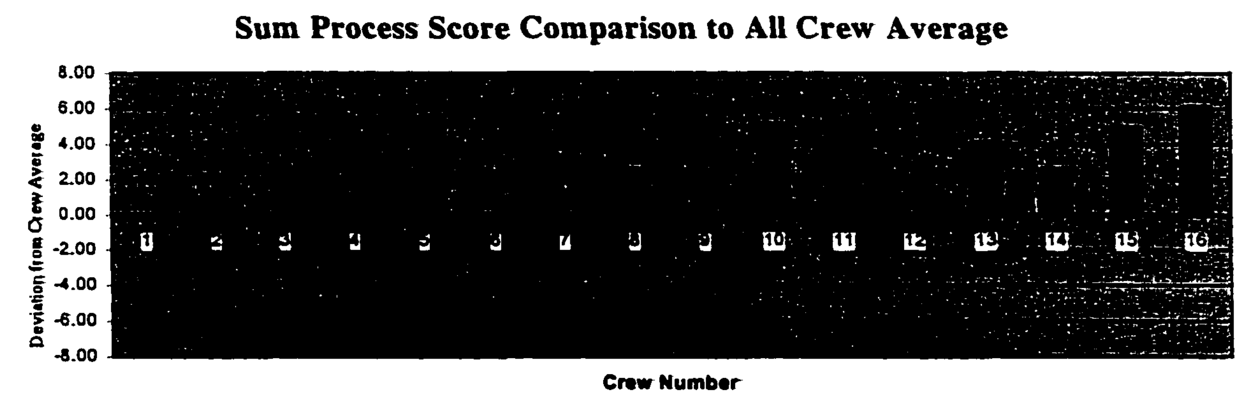

Figure 4. Comparison of individual crew sum process score to all crew average.

The mean differences in scores between novice and experienced crews for the seven process variables are reported in Table 6 and discussed below.

Table 6

Mean Difference of Process Scores Between Novice and Experienced Crews

Process Variable	Mean Score Difference
Take-off/Departure	.0313
Starter Caution Light for No. 1 Engine	.5125
Checklist Discipline	.3125
Low Oil on No. 2 Engine	1.2501
Approach Priorities	1.0000
Game Plan for No Arresting Gear	1.0750
Foul Deck	1.1875

A more comprehensive look at the component process ratings for each of the seven scenario events is shown in Figure 5. This figure corresponds to the level of granularity upon which statistical analyses were based. A significant main effect for experience on process ratings was obtained using a multivariate analysis of variance ($F = 5.974$, 7 df, $p = .011$). Given this overall statistically significant effect, the next question is to what extent do the processes within individual components contribute? A between subjects test summary table (Appendix I) reports the statistical significance of individual analyses of variance (ANOVA's) for each of the seven scenario events rated.

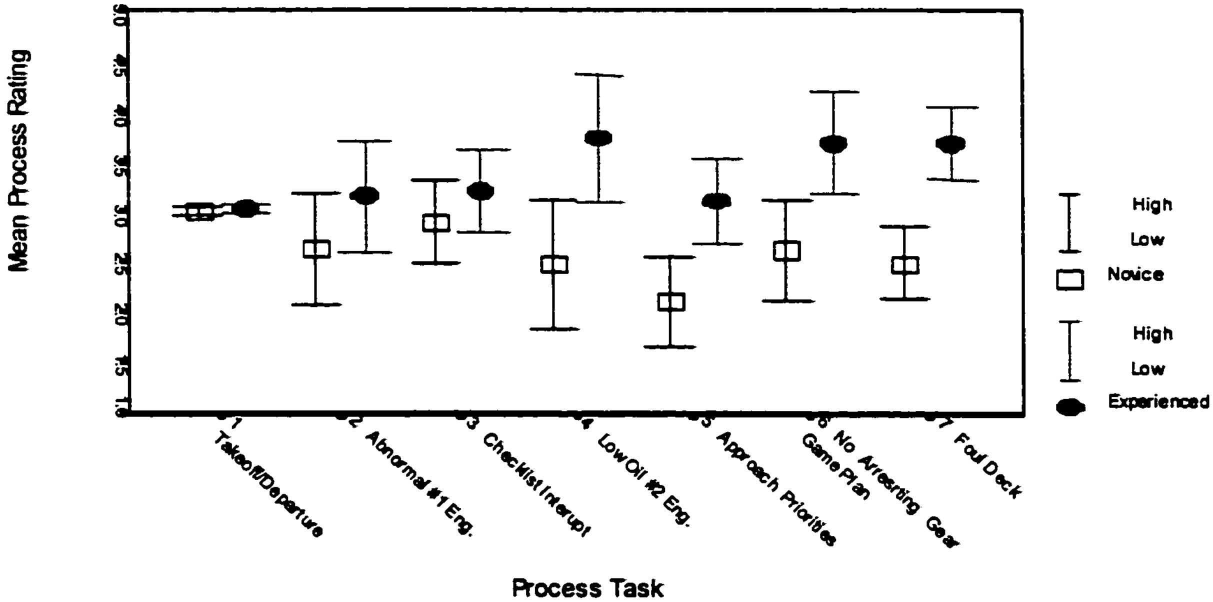

Figure 5. Mean process ratings of novice and experienced crews with confidence intervals.

These post-hoc ANOVA's revealed no statistically reliable effect for the simpler events (*takeoff/departure, Starter Caution Light illumination on No. 1 engine, interruption during checklists*) in the initial part of the scenario. Differences in means for the four remaining process events were significant using a 95% confidence level: (a) recognition and handling of *low oil pressure*, F = 8.61, 1 df; p =.011, (b) determining appropriate priorities, F= 12.10, 1 df; p =. 004, (c) *no arresting gear game plan* for landing a multi-engine aircraft with only a single properly functioning engine without the use of field arresting gear as required by Standard Operating Procedures (SOP), F = 10.39, 1 df; p = .006, and (d) making an "instantaneous" decision to continue landing or eject prior to or after the No. 2 engine failed as a result of low oil pressure, F = 24.12,1 df; p < .001. The results indicated that the means for the experienced groups' process scores for these four areas were significantly higher than the means for the novice group.

Aircrew Outcome Rankings

Hypothesis 2 predicted that the execution of the final decision results (outcome scores) of experienced aircrews would be higher than novice aircrews. As predicted, the superior ratings received by experienced crews, as a group, was statistically significant. The outcome data used by raters reflected the concluding "snapshot" of aircraft configuration and position relative to the landing environment. Raters' determination of most preferred aircraft and aircrew status was directly related to the aircrews' final decision to: (a) take the aircraft around again (in its degraded state) for another attempt to land, or (b) land immediately on the runway available and stop, or egress via an on-deck ejection prior to the stalled truck.

A one-tailed Wilcoxon matched-pairs signed-rank test was conducted to "evaluate the assumption that the two samples are randomly and independently drawn from similarly shaped populations with unknown but equal variation" (Berenson and Levine, p. 430, 1998). With the sum of outcome rankings of the novice crews being 45 and the sum rankings of the experienced crews being 91 the probability that these outcome performance scores would have naturally occurred is $p = .01$. This analysis shows that it is unlikely that seven of eight experienced crews would fall in the top 50% of the performance scores, while seven of eight would fall in bottom 50% if they had been drawn at random from a homogeneous population. Therefore, the superior ranking of experienced crews relative to novice crews was statistically significant.

Relationship between Process Scores and Outcome Ratings

The third study hypothesis predicted that aircrew process ratings would be positively correlated with the decision outcome ratings for a specific in-flight emergency

involving uncertainty was only partially supported with quantitative analysis. In Figure 6, the composite outcome scores are depicted on the vertical axis and the process ratings are summed across the seven scenario events and shown on the horizontal axis. Each symbol represents one of the sixteen crews, with solid circles representing experienced crews and open squares representing novice crews. The number above each symbol represents each crew's combined total S-3A/B flight hours. In this figure, the summed process ratings for experienced crews are generally clustered to the right of those from novice crews. The vertical axis represents outcome ranking scores with the rating of 1 corresponding to the lowest outcome score and the score of 16 representing the highest outcome score.

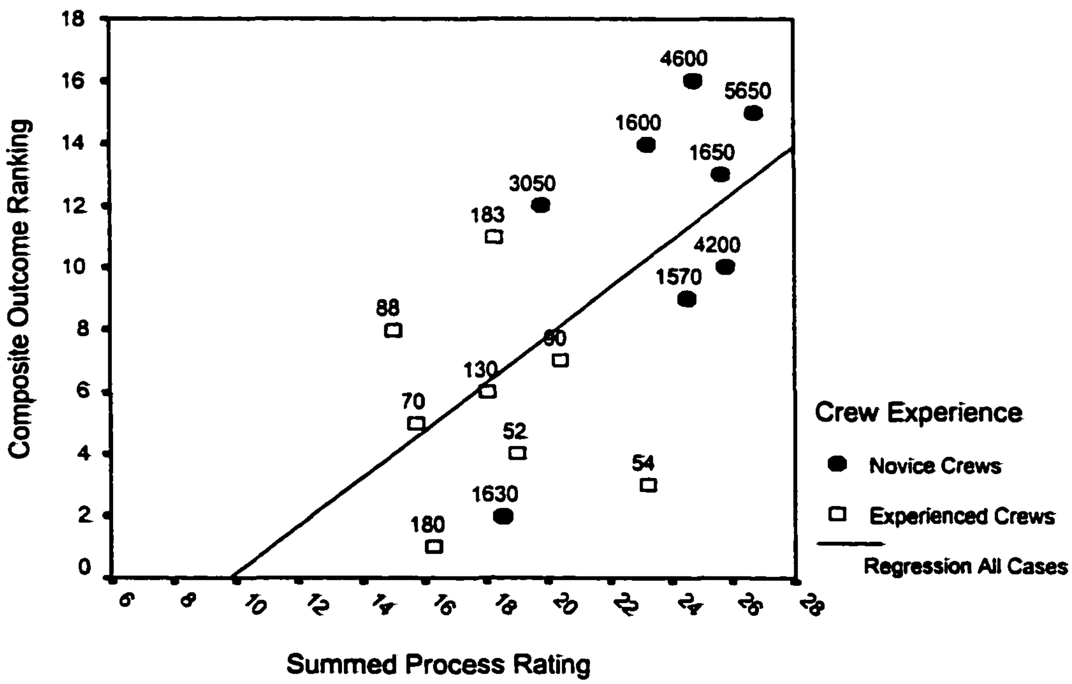

Figure 6. Relation between outcome ranking and summed process scores for all crews.

The mean rank order for experienced crews was 11.375, and 5.625 for novice crews. Each crew's flight hours relative to their process and outcome rank order are reported in Table 7. Novice crews 1 and 8 show a relatively good performance within their peer group and experienced crews 9 and 12 had comparatively poor performance judged against their peers. As can be seen in Table 7, in all but two cases, higher process scores and better outcome rankings distinguish experienced crews from novice crews.

Table 7

Aircrews' Combined Flight Hour Rank Compared to Process and Outcome Ranks

Crew	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Flight Hr Rank	12	13	16	14	9	10	11	15	4	3	7	6	2	8	5	1
Process Rank	7	15	9	14	11	13	12	*6	8	2	5	10	4	*6	3	1
Outcome Rank	10	9	13	12	6	16	11	14	5	7	8	15	1	3	4	2

Note. An asterisk denotes that these two crews received identical summed process scores.

Figure 7 depicts the overall final process and outcome score rank order for all 16 crews by assigned crew number. In the right column the top half of the outcome rankings consist of seven of the eight experienced crews (crews 9-16) while 7 of 8 novice crews (crews 1-8) fell in the bottom half of the outcome score rankings. Crew 16 had the highest process performance while Crew 13 had the highest outcome performance. As Figure 7 illustrates, there was distinct movement between process and outcome ranks. Yet, six of the eight crews who scored in the top 50% of summed process scores remained in the top 50% of outcome rankings. Although four crews shifted between top

and bottom ratings, there was a fair amount of stability within the process and outcome rankings.

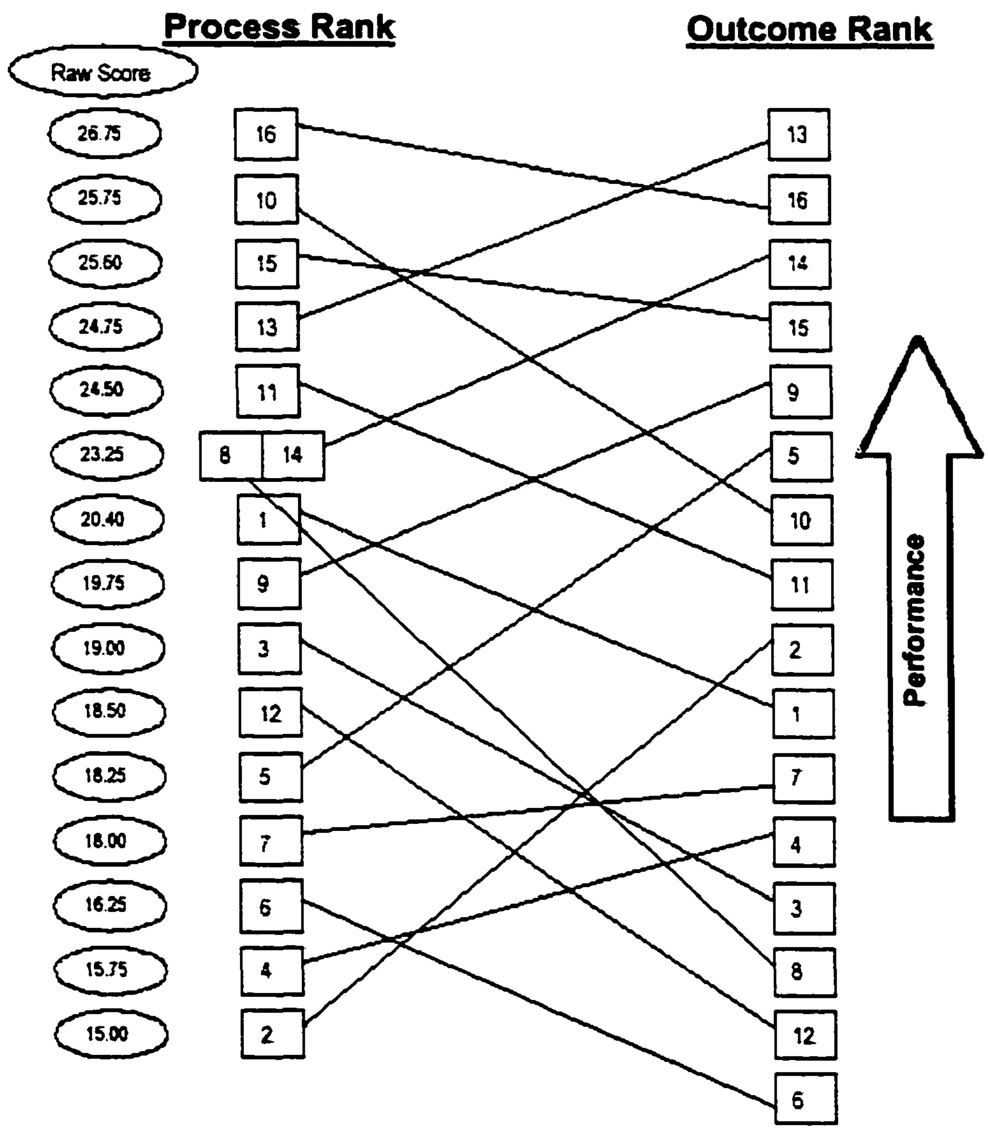

Figure 7. Comparative ranking change between summed process score rankings and final outcome rankings of all aircrews.

Analyses for the first two hypotheses clearly show that novice and expert crews represented two distinct populations with respect to both process and outcome. Due to

this lack of comparability, the experienced and novice groups were kept separate for correlation analyses. For novice crews, the process outcome correlation did not approach significance (r = 142, df = 7, *p* = .736).

A strong positive correlation (r = .64, *p* = .088, df = 7) was found within experienced crews, although the correlation was not statistically significant with the sample size used. Within the expert crews, notable positive correlations were observed between outcome rankings and more complex, unstructured decision contexts at the end of the scenario. Correlations between outcome and two process scores that approached significance were found in two events. These were *Foul Deck on Final* (r = .663, df = 7, *p* =. 073) and the *No Arresting Gear Game Plan* (r = .682, df = 7, *p* = .062). The correlation between process rating sum and the outcome (i.e., decision to eject) was also statistically significant (r = .695, *p* = .056, df = 7).

One of the more complex situations in terms of identifying and working with available information was recognition and handling the low oil pressure in the remaining engine although there was no statistically significant correlation of the time it took crews to recognize the drop in oil and ultimate outcome score. The mean time for novice crews to recognize a severe oil pressure drop in the No. 2 engine was 248.38 seconds (SD 201.14) compared to the experience crew mean time of 156.88 seconds (SD 180.82). Temporal data related to crew recognition of the low oil pressure is contained in Appendix J.

One of the objectives of this study was to identify and investigate issues related to sample size. A power analysis using a Fisher Z approximation was conducted to determine the sample size needed to detect a statistically significant effect "giving

consideration not only to the level of significance and the power of the test, but also to the effect size" (Hinkle, Wiersma, & Jurs, 1994, p. 312). If the correlation of .64 observed for experienced crews in this study reasonably approximates the correlation existing in the larger population, the probability of detecting a statistically significant correlation (p =. 05) would be .80 with a sample size of 16 crews.

The next section presents the findings of the three research questions associated with capturing the task specific thought protocols and performance skills of crews performing at three different levels (high, mid-range, and low) as defined by the quantitative analysis process and outcome rankings. Cross-case analysis findings include comparisons of process and outcome rankings with differences in decision strategies in uncertain, extremis circumstances requiring rapid situational and risk assessment.

Qualitative Cross-Case Analysis Findings

Selected naturalistic and aviation decision making theories and models guided the qualitative analysis. A variety of data sources describing participants real-time decision making in a dynamic environment were analyzed to answer the three research questions for the study. The three research questions sought the reasons for differing quantitative process and outcome findings of experienced versus low flight time crews challenged with identical in-flight emergencies involving uncertainty: (a) What adaptive recognitional/metacognitive decision-making strategies emerge from aircrews? (b) In what way do adaptive/metacognitive decision-making patterns differ among successful and less than successful aircrews? and (c) What characteristics/factors seem to define the most successful aircrew outcomes?

Emergent Adaptive Strategies

Analysis of data related to the first research question discovered that during the final three events in the flight, all crews attempted to adapt previous decisions to the shifting circumstances. Although crew effectiveness in adjusting to changing conditions and increased uncertainty differed. The crew's ability to meet performance requirements depended on collective experience, individual cognitive and technical skills, as well as their overall proficiency in crew coordination. Findings from the available data show that all crews' decision strategies were reflected in the scenario process performance scores and the final outcome rankings. It is significant that the ability to successfully recognize the need to adapt a strategic plan and follow through with an altered plan to successfully meet the requirements of a novel situation showed the highest degree of variation between experienced and novice crews.

Table 8 represents the range of decision-making and information processing constructs that were revealed during data analysis of the six crews representing high, mid-range and lowest scoring crews in process and outcome rankings. These decision strategies reflect a variety of methods employed by aircrews to attempt to control and/or transcend uncertainty during the flight scenario. Crews used these tactics to prioritize and reduce workload and increase focus under severe time constraints. As described in Table 8, these techniques assisted crews in identifying, sorting and managing limited and/or ambiguous information as well as assessing risk, and managing their workload. It is important to note that the success of a particular strategy was dependent on the timing and circumstances in which it was applied in the scenario.

Table 8

Adaptive Decision and Information Processing Strategies Detected in Study Aircrews

Decision-Making Strategies	
Uncertainty Circumnavigated	Crewmember(s) devises a game plan that completely negates their need to deal with the uncertainty
Uncertainty Eliminated	Crewmember(s) adopts a tactic of "forced resolution" and takes deliberate action intended to resolve the uncertainty by eliminating other options as clarification/solutions
Uncertainty Acceptance (or) Uncertainty Carried Forward	Crewmember(s) recognizes and identifies elements containing uncertainty factors yet chooses to press ahead regardless
Uncertainty Ignored	Crewmember(s) acknowledges uncertainty factors and then simply ignores them and presses ahead with a "blinders on" mentality
Situation Acceptance	Crew member(s) accepts situation with no attempt to resolve uncertainty or plan around it
Activity Acceleration	Crewmember(s) accelerates the completion of routine administrative items in anticipation of upcoming periods of high workloads or entering into more demanding or dynamic environments
Information Seeking and Use	
Info Gathering Expedition	Crewmember(s) sets out to find additional cues to help resolve uncertainty (either visually: instruments, sight picture, control positioning or verbally: asking questions of other crewmembers and or external resources)
Information *Firewalling*	Crewmember(s) intentionally delays or ignores (compartmentalizes) reception or introduction of new information (Relative to external sources: crew denies the information source the opportunity to communicate via the radio or some other means, i.e. "stand-by")
Assimilation Avoidance	Crewmember(s) acknowledge the presence and availability of new or additional information but simply chooses to "leave it alone" and not process or act on it
Curiosity Flat line	No attempt to gather additional information

Patterns of Decision Making and their Relation to Performance Levels

In general, the findings related to the second research question established that aircrew experience level strongly influenced strategies employed to successfully meet performance requirements under different naturalistic circumstances. The more successful, high flight hour crewmembers more adeptly processed and shared information given interrupted routines and insufficient time to support their original strategic goals. The ability to recognize cues and similarities in patterns, to rapidly assess dynamic and/or novel situations, and to make the necessary adjustments and achieve a successful result were demonstratively related to the crews' domain knowledge, experience level, and team skills.

These findings support the relationship between best possible results (i.e., outcome ranking) in the study scenario and the successful use of metacognitive skills to continuously self-monitor, critique, and correct thinking strategies to assess and project a "simulated" course of action in a situational model to achieve optimum results under novel, dynamic conditions. This adaptive decision making process applied by individuals to meet severe decision making requirements was framed and described in the recognitional primed decision (RPD) model (Klein, Calderwood, & Clinton-Cirocco, 1985; Klein, 1989) and related studies. These studies have shown that under dynamic conditions of the naturalistic decision environment, experts use a more intuitive approach to meet the demands of rapid troubleshooting and mental simulation to select the first reasonable course of action that will satisfy immediate problem requirements. The optimal cognitive processes used to rapidly gather and assess relevant information for accurate situational assessment and decision making both consider and surpass the procedural or more analytical approaches. The routine procedural and analytical decision

making strategies either do not meet the situation requirements or use more time than is appropriate (or available) for the circumstances.

Key functions of recognition/metarecognition processes. The recognition /metacognition (R/M) model complements the RPD model by addressing the metacognitive aspects of the situational assessment and decision-making process. As defined by Cohen et al. (1996) meta-recognitional processes "determine when it is worthwhile to think more about a problem; identify evidence-conclusion relationships within a situational model; critique situational models for incompleteness, conflict, and unreliability; and prompt collection or retrieval of new information" (p. 206).

Participant retrospection of their conscious thinking processes fell into functional areas associated with adaptive decision making. Time consuming, concurrent option weighing to achieve an optimum solution was replaced by selection of the first acceptable sequential option. Data from crew debriefs were analyzed for "fit" into the three functional areas of the meta-recognitional cycle: (a) critiquing or accurately evaluating/characterizing the problem, (b) monitoring a course of action to assess whether the methods and results of the decision process will be satisfactory, and then (c) correcting or regulating the plan with a sequential evaluation of options with a commitment to the first acceptable alternative rather than trying to optimize by waiting for analytical results (Cohen et al.,1996).

Aircrew self-reports of strategies and process content. The study findings support that the ability of the aircrew as a whole, not simply individuals in the aircrew, to adapt to the cognitive requirements for decision tasks in each phase of the recognitional cycle is crucial to using adaptation strategies successfully. Because the R/M model goes beyond

the processes used in situational assessment, it is valuable in its use here to frame and describe the aircrews' meta-recognitional cyclical processes in dealing with uncertain and novel situations in the three distinct areas described above. This model was adapted to assist in discovering patterns in team performance by overlaying the cockpit-specific analytical scheme developed by Orasanu, Dismukes, & Fischer (1993) to predict types of errors based on different cognitive requirements for various decision making situations in a multi-crew cockpit. A sample of participant quotes that relate to situational assessment, determination of a game plan, and crew coordination issues were integrated with the adaptive decision making described in the R/M cycle described above and analyzed within the cockpit decision error framework. These data allowed comparison of types of decision-making strategy used (analytical, option-based, and adaptive/creative) as well as comparison and analysis of process errors in the details of content of the metacognitive process reported by the aviators during the debrief/interviews.

Examples of pilot and COTAC (navigator and co-tactical officer) recall of metacognitive activity are presented in the following succession of quotes and are characterized as they relate to the three functional areas of the meta-recognitional cycle described above. As a reminder, crews 1 through 8 are novices and crews 9 through 16 are fleet experienced crews. Note that Crew 12, an experienced crew, was ranked 15 of 16 in the final outcome ranking and that Crew 5, a novice crew was ranked 6 of 16 in the outcome ranking.

In general, the content of metacognitive thoughts of experienced crewmembers as quoted below reflect that they were frequently searching out information and verifying their situational assessments. The starter caution light and low oil pressure events

provided opportunities to observe crews as they worked through the information gathering and decision-making cycle by recognizing and verifying cues, working together to make certain that cues were being interpreted correctly, assessing resources, and setting priorities within time constraints. However, the debrief/interview provided an opportunity to gain greater insight into the metacognitive/meta-recognitional process as relayed by experienced and novice aircrews. Samples of these tacit processes are provided starting with the first phase of the R/M cycle.

Critiquing or accurately evaluating/characterizing the problem. The pilot of Crew 13 as well as the COTAC of Crew 9 is evaluating the initial cues associated with a problem while facing a deteriorating situation. They continue to use their critical information seeking and use skills to further evaluate the problem to look for and ensure secondary indications are correctly interpreted and evaluated.

"I think once we noticed the oil pressure dropping in the second gage then at that point in the decision [making process] and once we saw that that's [field arresting gear] not engaged then it's get the plane on deck as soon as possible and whatever field [amount of available runway] we land with we land with." (Pilot, Crew 13)

"The first thing you do when you shut down the engine is to make sure all the lights [on the advisory panel] that you get correspond to what you expect to see. It was also the one light that can trick you [the Engine Oil Pressure light does not reset to re-illuminate if there is subsequent loss of oil pressure on the remaining engine]..." (COTAC, Crew 9)

The quote below from the Crew 12 pilot (crew 12 ranked 10 of 19 in process and 15 of 16 in outcome rank) provides insight into several of the underlying issues related to this crew's poor performance. The pilot misjudged the level of risk based on a lack of knowledge, his oversimplification of the problem, and his neglecting to consider viable options. He treated the option to re-start the No. 1 engine as a rule-based decision. By not moving forward and accepting a calculated risk associated with this somewhat

unorthodox procedure (restarting a previously secured engine) he accepted the uncertainty in the situation (time remaining for the No. 2 engine until it stopped functioning) and lost the opportunity to generate a potential back-up option in case the No. 2 engine failed (due to loss of oil pressure) prior to reaching a position from which either a safe landing or a successful ejection could be accomplished.

> "I set myself up to start up the No.1 at the beginning indications that the starter was having problems. [Pilot seems to indicate that once he recognized the possibility of losing the No. 2 engine [due to low oil pressure] he configured the No. 1 engine switches so that the crew would be ready to attempt a no. 1 engine re-start rapidly.] So we had ourselves in the position where all I had to do was starter switch "engage". At least I hoped so. So I thought we were ready to clean up on that. I didn't see the need to start it though. I sure wasn't going to compound the emergency." (Pilot, Crew 12)

Examples of uncoordinated task performance are described in the two quotes from the novice pilot in Crew 2 that follow. Although the pilot was able to retrieve procedurally prescribed responses normally associated with a failing engine, he did not keep his COTAC in the information loop. Later in the flight, still uncertain about the reliability of the No. 2 engine, the pilot did not take into consideration the coordinated tasking required to restart the No. 1 engine while simultaneously flying the final approach. Even under normal circumstances the final approach is a high workload sequence of events.

> "... with one fluctuation and a little bit after I had to figure out a few things. We did have an engine vibration problem. With the simulator it makes it really hard to tell that so I was thinking we just did that precautionary aspect treating it that way because I was already at IDLE. [Indications associated with both engine vibrations and low oil pressure situations require the throttle to be retarded to the IDLE position.] You know, slowing down my descent, cycle bleed just in case that was it. That or T-5 thing [referring to an instructor introduced simulator malfunction designed to draw the crew to scan near the oil gage.] I didn't see an

issue there because I wasn't revving up on it [the remaining engine]. The next aspect was pending engine failure. I mention that because I don't know how well I verbalized it." (Pilot, Crew 2)

"I'd rather risk screwing up an engine [identification of potential problems associated with restarting the No. 1 engine with a faulty starter indication. The implication being that the pilot is willing to risk damaging or setting the No. 1 engine on fire if it provides a better chance to save the jet]. All I have to do is shut it down because it's going to go bad. I'd rather risk that engine than risk losing the No. 2 engine and forcing an ejection. Once I bought that turn right there –the turn away from the runway...that's it. [Concern he has enough altitude and airspeed to *only* make a single turn. If that turn is away from the airport then he'll be committed to ejecting from the aircraft over the water.] We have a certain amount of minutes before we can make it around again, turn to 29. Okay? Well, we're doing that right now—we're doing that already, we've bought [committed to] that already so we have time to breakout the Checklist, if I say so, and try to start No. 1 [engine]. We've already bought [committed to] what we want. That's the logic of it- I mean, trying to start No. 1". (Pilot, Crew 2)

Monitoring a course of action. As the crewmembers' quotes below explain, the experienced COTAC of Crew 11 continued to monitor and adjust the plan of action to keep the crew's focus on making an expedited landing. The pilot of Crew 14 also displayed assertive creative problem solving skills as he considered the requirements to land in a situation that did not allow the optimal time to complete checklists and to evaluate and discuss multiple options and contingency plans.

"At that point, when I told you [pilot] to go ahead and don't worry about the gear speed limitation [speeds in excess of the prescribed speed limit run the risk of damaging either the aircraft or the field arresting gear equipment, or both] I was actually just monitoring our progress and the amount of time that the engine had left. I was very much interested in getting the airplane back to the runway." (COTAC, Crew 11)

"I think that [truck in intersection] was probably the hardest decision... the runway is semi-clobbered...Do you still want to continue with this knowing there is no short field gear? [Both NATOPS and SOP's require arrested landings in single engine landing configurations.] I think we paused for a second and then continued with it. (Pilot, Crew 14)

In contrast to the aviators in crews 11 and 14, COTAC's in crews 15 and 12 did not interpret the situation correctly which led to their neglecting relevant information and delaying the completion of essential tasks. The inappropriate conclusion by the COTAC in Crew 15 indicates that even after he was aware of the low oil pressure indication (an obvious indication of the rapidly deteriorating condition of the only remaining engine), he lacked sufficient situational awareness to grasp the "big picture" and recognize the severity of the situation. In this scenario, attempting to divert to another field over populated areas was not a viable option due to: (a) the diminished amount of thrust that the No. 2 engine was producing, and (b) the highly questionable amount of time that the engine could be expected to continue to operate.

Crew 12 was the only crew not to identify the low oil pressure on the No. 2 engine as indicated by the No. 2 Engine Oil Pressure gage. When the control tower informed them of white smoke trailing from their starboard engine this crew chose not to engage in any additional troubleshooting or information gathering activities. Instead, they simply acknowledged that the white smoke coming from the engine signaled that there was a problem that complicated their situation. (In this case, the lack of awareness of the cause of the problem with the No. 2 engine resulted in the pilot executing a wave-off without associating throttle advancement with engine failure.) The Crew 12 COTAC decided that getting the checklists completed should be the priority and therefore he attempted to delay the approach and landing in order to complete his portion of the routine cockpit activities. This COTAC oversimplified the problem and failed to recognize that any of the prescribed procedural responses found in NATOPS (the aircraft operating manual) did not account for the circumstances he was facing. Without

adjustments to expedite a normal approach profile (and associated COTAC routines) there was a very high likelihood that the single remaining engine would quit or fail in the time taken to complete the prescribed "normal" procedures.

> "We were below "max trap" [maximum aircraft fuel weight allowed to make an arrested landing] when I figured dumped fuel is no good to us in case, for some reason, we need it back at some point even though we had made the decision to land and I know we're below max trap. More fuel is better than no fuel." (COTAC, Crew 15)

> "... I didn't verbalize it, but I thought briefly about just setting ourselves up for the final so we could just go land instead of having to deal with the DELTA pattern [orbiting overhead the field in a prescribed flight path]. But then I was like, 'I know we're single engine and I know we've got at least two more Checklists to do. So, yea, let's just go ahead and orbit in the DELTA pattern." (COTAC, Crew 12)

> "So at that point, I'm thinking, hey, we can still make a normal single-engine landing. The nosewheel steering was functioning. We did eat the arrested landing. That was just precautionary and I was thinking, Let's go ahead and continue per NATOPS and just land."(COTAC, Crew 12)

Correcting or regulating a plan. The following participant quotes are examples of experienced crews describing the retrieval and review of situational constraints as well as their generation and evaluation of options that led to successful outcomes in the final scenario task.

> " I was [high] on purpose because I wanted to keep power back to try to save the engine as much as I could in case we needed to use it and then I was like, Well, we're not going to go around [wave-off] unless something really goofy happens. So when the Tower said [there was a truck in the runway] at that point I wasn't thinking I wasn't going to go around unless I was actually going to land on top of something. I just wanted to land at the end of the runway then when you told us all those things. So, instead of landing a normal trap [field arrestment] and landing by the gears [field arresting gear], land at the end slow [below normal approach speed] and just get on the brakes and try to stop." (Pilot, Crew 10)

> "... one of the things that I was thinking about that if it's real -- if the runway is completely clobbered and we're going to need the distance we could go to the taxiway and land on that. But that never really became an issue. I briefly toyed

> with that idea and then disregarded it for landing on the runway and try and get it stopped." (COTAC, Crew 11)
>
> "Just from experience you know you can pretty much normally stop by that point [the intersection where the truck was located] and if we couldn't it would be a fairly low speed taxi clear, maybe depart the runway at a low speed, or maybe a high speed turn off into a taxiway. It was what was going on through my mind." (Pilot, Crew 14)

As typical of many of the novice crews, Crew 1 pilot appears to recognize the urgency of the situation and makes adjustments to reduce the routine number of checklists down to just the Landing Checklist. However, he does not project what he will need to do to get in a "good landing position" and fails to precisely monitor or correct his approach to land. He lacks the forethought and technical skills to affect his desired outcome in the time available. This pilot terminated his first approach attempt to make the runway, used valuable time to come around again and then executed a wave-off on his second approach.

> "Shut Down the No. 1 engine –that's obligatory. Get plane on deck after low oil recce [recognition] in No. 2 [engine], that's when I knew that everything, all the checklists we really need to do, all we really needed to do was just get the plane on deck. Really it was just the Landing Checklist. We need to get in a good landing position." (Pilot, Crew 1)

As illustrated by the preceding quotes from aircrews in the three stages of the R/M cycle, it appeared that although novice groups were working through portions of the cycle associated with adaptive decision making, they were not identifying or using the higher levels of cognitive work required to meet the requirements of the decision tasks that were facing them. The crews that choose to wave-off were unable to adapt their decision strategy to correct and adjust their plan for the best possible outcome. As a

result, the lower performing crews relied on more familiar procedural and analytical based solutions to a novel problem that required an adaptive strategy for the best possible outcome. Beyond the selection and use of an appropriate decision making strategy to meet unique situational requirements, there are also crew attributes and crew functioning characteristics that represent distinctions that separate process performance levels and outcomes. In this scenario crew coordination between crewmembers and effective levels of crew communication were crucial in the last minutes of the flight to correct and refine the plan of action.

Novice Crew 6, in particular, provides one of the clearest illustrations of the importance of crew coordination as part of the joint situational assessment and decision-making cycle in a multi-crew aircraft. This crew's performance also illuminates the necessity for maintaining a high degree of crew coordination throughout the flight. Crew 6 was ranked 14 of 16 in process performance and was ranked last in the outcome of their final decision in the scenario. One of the significant problems in this crew was the pilot and COTAC had different objectives that led each one to execute contrary game plans. The pilot had made a determination to eject while the COTAC was still focusing on continuing the landing. These divergent motivations caused a fair amount of crew disconnection in the final seconds prior to the crew abandoning the aircraft. Appendix K contains significant dialog of this crew's interactions during the last portion of the flight.

Decision Making Process Pattern Differences among Case Studies at High, Mid-Range and Low Performance Levels

The second research question of interest was whether there were differences in decision strategies between crew process performance levels. Another issue investigated was whether strategy differences correlated with better or worse outcome ratings. In order to explore these issues, process related data and metaperceptual data from three crews with the highest (16), mid-range (3), and low (2) performance levels were analyzed and sorted by strategies associated with situational awareness and selecting a course of action.

The qualitative analysis results mirrored the quantitative findings: that the most prevalent differences in strategy selection occurred in the most novel decision context (i.e., whether to land or wave-off to attempt another landing). The qualitative investigation also found that there were distinct differences in the approaches of crews 16, 3, and 2 for resolving the uncertainty in all the increasingly unstructured events. Figure 8 represents a comparative overview of the specific procedural, analytical and adaptive decision strategies employed by these three crews during the final scenario events. The larger font indicates the type of strategy that was most pronounced during that phase of the flight. The low oil pressure indication and approach events have been combined. These two events were combined in the comparative analysis of crew decision making because the recognition of the low oil pressure condition occurred at different points in the approach for each of these crews. The three event sets depicted in

Figure 8 are discussed in detail in the next section of this chapter.

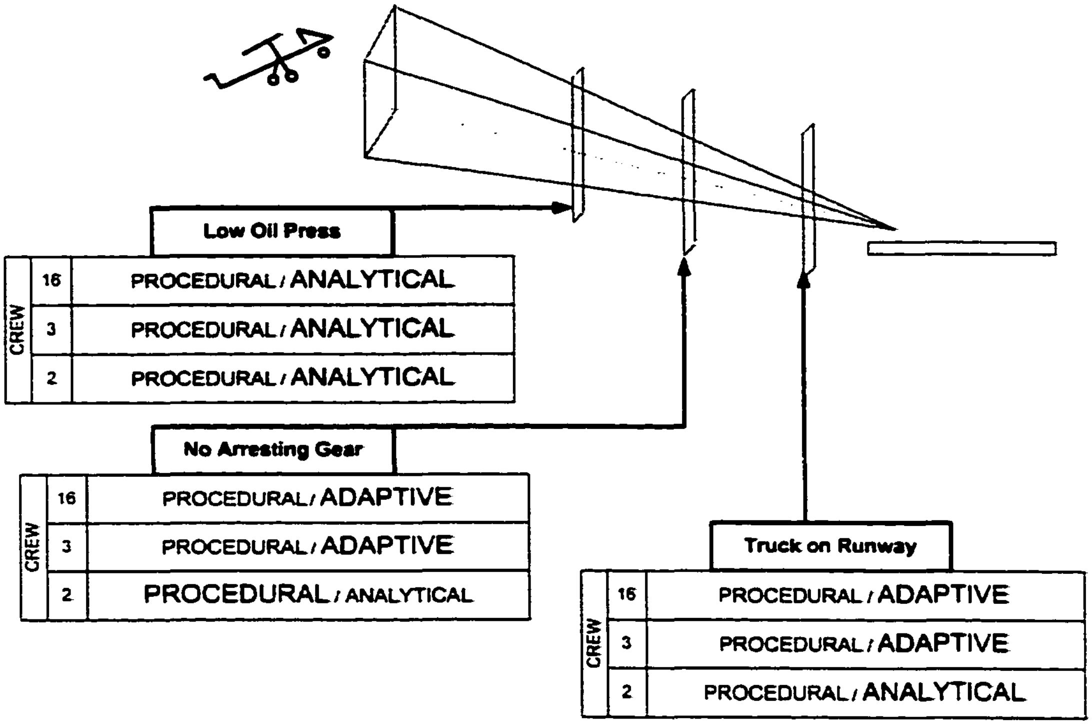

Figure 8. Decision strategies employed by high, mid-range and low process ranking crews (crews 16, 3 and 2)

Tables 9 through 11 present a linear "snapshot" description of the cyclical crew procedural, analytical and/or adaptive decision making process represented in Figure 8. Aircrew activities are categorized by participants' metacognitive self-reports and observed actions related to the final sequence of significant events in the study scenario.

In-Depth Case Analysis

Table 9 records the data related to the crews' overt and metacognitive processes related to the recognition and reactions to low oil pressure in the No. 2 engine (i.e., only remaining engine). During the low oil pressure and approach event, all three crews discuss and weight options at this decision point in the scenario; yet there is a marked difference in the resulting focus and game plan developed by each crew. Since there is

no distinct adaptive strategy used at this point in the scenario, Table 9 reflects only procedural and analytical categories. The more experienced, higher performing crew (16) is clearly focused and able to proceed with the approach to land since they have accomplished/coordinated the completion of aircraft configuration changes, checklist routines, contingency planning, and landing preparations thus far in the flight. Unfinished cockpit routines and lack of communication and cooperation in the novice crews starts to catch up to both Crews 3 and 2.

Table 9

Specific Examples of Strategies Used by Aircrews under Conditions of Uncertainty and Increasing Time Constraints during Low Oil/Approach Events

Strategies Used	**High Performance (Crew 16)**	**Mid-Range Performance (Crew 3)**	**Low Performance (Crew 2)**
Procedural			
SOP/Checklists	Fuel Dump Completed in transit. All Checklists Completed except no Before Air Start. Checklist for No. 1 re-light attempt discarded for sake of expediency.	Trade-off fuel dump for checklist completion. Incomplete Landing Checklist/No Wave-off Brief	Did not complete SE Checklist. No Approach Checklist. Wave-off Brief conducted at 30 miles out. Orbited to dump fuel at 20 miles out.
Analytical/ Weighing of Options			
Situational Awareness Activity	COTAC recognizes low oil pressure. Pilot checks aircraft position heading for populated area	With field in sight (4.3 miles) Pilot requests to DELTA overhead then recognizes low oil pressure	No recognition of low oil from gage or scan of engine tape fluctuation. Tower calls to inform of smoke from No. 2 engine **(Table continues)**

Strategies Used	High Performance (Crew 16)	Mid-Range Performance (Crew 3)	Low Performance (Crew 2)
Info (cue) use	Pilot," Give consideration to No. 1 restart." Pilot changes flight path on approach to avoid populated area in case of ejection. Still confident he can land. COTAC, "Things are not looking good."	Does not report low oil to COTAC until transmission from tower about white smoke from No. 2.	Crew confused. Asks for clarification on whether smoke is from wing or engine. Pilot associates smoke with low oil and takes radios. Tells ATC to cancel IFR (Instrument Flight Rules) and proceed direct to field (visually) for arrested landing. Throttle to IDLE.
Weigh/Discuss Options and tradeoffs to create a plan	Continue on shortest route to field. Pilot talks through options related to restart. COTAC (meta) Takes a while for starter to degenerate (from own experience).	Pilot aware of No. 2 low oil pressure. Wants to get down ASAP for arrested landing on 29 but may have to shut down #2 and eject Tells COTAC to "start in on those checklists."	----
Plan accepted?	Yes. Pilot to Tower, "we're going to do a full stop landing." COTAC (metacog) – weights probabilities of fate if different runways missed	Yes	Yes. COTAC tells tower "we're going to bring it down and remain on deck." (table continues)
Contingency Plan?	If can't make field to land on numbers- Eject over water	Eject	COTAC to Pilot "I recommend we don't do a wave-off." COTAC (meta) "keeping fingers crossed hoping No.2 didn't crap out."
Attempt to gather more info?	No discussion	COTAC dismisses Tower call about white smoke from starboard (#2) engine as Tower's probable confusion with No. 1 engine.	No discussion
Operational pace	Efficient	Expedited Landing	Slow

Note. A dashed line (----) denotes no activity observed or reported by the crew.

Data in Table 10 represent actions and thoughts from the same three crews while they are flying with one engine of extremely questionable reliability on their final approach. The crews were informed that there was no arresting gear on the only available runway. Both NATOPS and standard operating procedures direct crews conducting single-engine approaches to make arrested landings. An arrested landing allows the aircraft's tail hook to engage the arresting cable strung across the runway. Engaging this cable drastically reduces the aircraft's landing roll out and eliminates the need to use aircraft systems that may be degraded or taken "off-line".

At this point in the flight Crew 16 generated a plan that forced resolution of the uncertainty issues. A vigilant execution of a plan to achieve a strategic goal (i.e., "let's land") with expert technical skills enabled Crew 16 to fly a precise final approach. The crew achieved their goal by focusing their cognitive resources and eliminating irrelevant issues and information. In other words, they decided to press ahead and land without diverting their attention to non-priority issues. In actuality, this is an elegant and extremely pragmatic solution to an exceedingly complex problem. This crew determined that all their engine problems would become inconsequential (or irrelevant) if they simply landed and stopped the jet on the available runway. Crew 16 provides an excellent example of an adaptive decision strategy called *satisficing* (Simon, 1955). Satisficing is an approach "for making a choice from a set of alternatives encountered sequentially when one does not know much about the possibilities ahead of time . . . there may be no optimum solution for when to stop searching for further alternatives . . . " (Gigerenzer & Todd, 1999, p. 13). Crew 16 successfully uses a recognition process that obviates the

need for further information and adopts a realistic option to land immediately and foregoes an attempt to generate an optimum solution.

Table 10

Specific Examples of Strategies Used by Aircrews under Conditions of Uncertainty and Critical Time Constraints during Final Approach with No Arresting Gear

Strategies Used	High Performance (Crew 16)	Mid-Range Performance (Crew 3)	Low Performance (Crew 2)
Procedural			
SOP/Checklists	All Checklists completed.	Did not complete numerous checklists (Approach, Single Engine, Arrested Landing). Landing Gear still up.	Pilot initiates Landing Checklist at 3 miles. Does not reset Fire Pull Handle in No. 1 re-light attempt (from memory).
Analytical/Weighing Options			
Operational Pace	Quick/efficient	Pilot senses immediacy but does not relay to COTAC	Rushed-pilot attempts restart from memory. COTAC nonchalant
Situational Awareness Activity	----	----	Pilot too high on first attempt- goes around for landing
Info (cue) use	----	----	Does not stress No. 2 engine. Utilizes time to attempt No. 1 re-light
Weigh/Discuss Options and tradeoffs to create a plan	----	----	Pilot wants to attempt to Re-start No.1 at 800 ft on 3-mile approach (after being talked out of it earlier by COTAC). COTAC wants to finish landing checklist Pilot wants to troubleshoot. COTAC "We're opening a whole new can of worms."
Plan accepted?	----	----	COTAC uncooperative- refuses to backup Pilot with Checklists **(Table continues)**

Strategies Used	High Performance (Crew 16)	Mid-Range Performance (Crew 3)	Low Performance (Crew 2)
Contingency Plan?	----	-----	COTAC to Tower-"We're going to bring it down... and remain on deck if we're not going to stop by the end of the runway we'll get out [eject]. All right?"
Adaptive/Satisficing			
Operational pace	Expedited Approach.	Sense of urgency by pilot but not COTAC until Pilot shares info about No. 2 engine status.	----
Situational Awareness Activity	----	Pilot informs COTAC he has "30 seconds to complete checklists".	----
Info (cue) Use	Pilot more "comfortable" with Runway 29 than 36	COTAC responds, " That No.2 engine?"(Still unaware of No. 2 engine status) Land if no ejection required prior	----
Attempt to gather/delay more info?	Pilot actively delays further info from Tower	Tower told to "Stand- by"	----
Accept risk-Press ahead with focused plan	Pilot (metacognition) "The only thing I wanted was an arrested landing."	Pilot (metacognition) doubted his ability to stop aircraft	----
No Contingency Plan. Accept situation with no attempt to act to effect probable outcome	----	COTAC informs pilot he's at 900 ft with 2.5-descent rate. Pilot responds," We have to get down with the airplane." Pilot: "We've got 15 psi-on that right engine so we're going to be getting out of this jet or we're going to land."	----
Contingency Plan?	No	No	----
Plan accepted?	Yes	Yes	----

Note. A dashed line (----) denotes no activity observed or reported by the crew.

Crews 2 and 3 are still dealing with issues of uncertainty resulting from a lack of information, poor communications, and less than optimum task management which are interfering with their ability to mentally "keep ahead" of the aircraft. The Crew 3 pilot has not communicated the low oil pressure problem to his COTAC. Now the decision making process is slowed by the COTAC's need to re-sort and recognize internal and external communications about uncertainties in information and time constraints.

Crew 2 is dysfunctional in terms of their ability to achieve consensus internal to the cockpit, let alone coordinate with the air traffic controllers to talk about further flight clearances. The pilot is attempting to raise the probability of sustaining controlled flight by attempting an in-flight restart of the No. 1 engine. The COTAC simply accepts the uncertainty of the situation (i.e., the possibility that the No. 2 engine may fail at any time) and does not want the additional tasking (consulting the published checklists) affecting a potentially better outcome. The COTAC asserts that the probabilities of needing the back-up engine are outweighed by the possibility of a restart "explosion" over a populated area. The pilot does not assert his positional leadership to order the COTAC to initiate the Checklist "challenge and reply" routines. The pilot attempts to re-light the No. 1 engine by himself from memory without the benefit of the written procedures (available only to the COTAC in the checklist). The timing of this request was rejected by the COTAC as more work than the COTAC could deal with at this stage of the flight relative to the impending final approach and landing, the aircraft's mechanical state and his own lack of "mental reserves."

Although there are attempts at adaptive strategies, novice crews 2 and 3 take too much time analyzing and creating options to effectively deal with the situation. They

appear to be waiting for "someone" to provide them with direction or guidance (in this case the airport tower) or for something to happen (perhaps the failure of the No. 2 engine) to force them into a decision. Crew 3 attempts to increase the pace of the cockpit routines, as they perceive scant minutes available for them to complete the remaining non-essential for safety of flight procedural tasks normally associated with a landing sequence. However, they stay in the analytical mode far too long while they attend to unresolved issues concerning aircraft status. Similarly, Crew 2 COTAC does not adapt to the situation; he spends valuable time relaying two plans to the control tower instead of backing up the pilot.

In comparison, high performing crew 16 continues at an efficient pace, employing good technical skills for the approach and good information exchange in the cockpit. There is no discussion of contingencies. In contrast to the other crews, the pilot of Crew 16 begins to focus the entirety of his attention on landing and adopts a mindset that filters out or eliminates all non-essential (i.e., non-landing) stimuli and influences. He commits full concentration to completing an "arrested landing." The COTAC provides aggressive pilot backup by minimizing external communications during this critical phase of flight.

In the last most difficult and time constrained decision in the scenario the crews are faced with a novel, dangerous situation that they must resolve while continuing to sustain the high workloads, sustain potential energy to reach the runway, scan for other aircraft, and configure the aircraft for landing. The aircrews' strategies to deal with a compounded novel set of high-risk conditions in a dynamic environment are outlined in Table 11 that follows.

Table 11

Specific Examples of Strategies Used by Aircrews under Conditions of Uncertainty and Severe Time Constraints during Final Approach Tower call about Truck at Runway Intersection

Strategies Used	High Performance Crew (16)	Mid-Range Performance Crew (3)	Low Performance Crew (2)
Procedural			
SOP/Checklists	Speed brakes employed. Flaps in Take-off position Hook down.	----	----
Analytical/Weighing Options			
Operational pace	Rapid	Rushed	Measured
Situational Awareness Activity	----	----	----
Info (cue) use	----	----	Pilot assumes entire runway fouled.
Attempt to gather more info?	----	----	----
Weight/Discuss Options and tradeoffs to create a plan	----	----	COTAC (meta) "How am I suppose to tell a pilot what to do?" (Previously crew determined "No wave-off.")
Plan accepted?	----	----	Now plan accepted. Pilot waves off and crew ejects over runway @ 50 feet.
Contingency Plan?	----	----	Yes. Several if they had landed. **(Table continues)**

Strategies Used	High Performance Crew (16)	Mid-Range Performance Crew (3)	Low Performance Crew (2)
Adaptive/Satisficing			
Operational pace	Efficient but hurried	Urgent	----
Situational Awareness Activity	----	COTAC makes altitude calls. Pilot unsure of truck location (in middle of runway or intersection or end of runway). Knew they were heavy	----
Info (cue) Use	----	Pilot does not verbally respond. COTAC (meta) "Pilot fighting the jet- maybe he didn't realize how low he was." Pilot unsure of ability to stop aircraft before truck. "We would have probably hit it." Pilot knew putting throttle to firewall would not give them another attempt at approach. Probability high for ejection	----
Attempt to gather/ delay more info?	Pilot ignored/missed call COTAC acknowledges transmission and tells tower to "Standby"	No	----
Accept risk- Press ahead with focused plan	Crew lands on numbers and rolls to stop before truck.	----	----
No Contingency Plan. Accept situation with no attempt to act to effect best probable outcome	----	Pilot expected to eject. Pilot to control tower, "Roger, we're just going around."	----
Plan accepted?	----	Yes. Ejected over water at 100 feet. However, COTAC tried to get call out after pilot calls for ejection over water. COTAC is late to pull handle (attempts 2 times) to eject to ensure search and rescue (SAR) effort gets under way. C- Standby 701 is ejecting P- Eject, eject, eject C- Ready?	----

Note. A dashed line (----) denotes no activity observed or reported by the crew.

The strategies used by both Crews 2 and 3 to circumnavigate the uncertainty of the truck position on the runway and/or the ability to stop or eject prior to the truck led them to wave-off on the final approach. The wave-off required full power to be applied to the only remaining, but already degraded and poorly performing engine. This application of power resulted in an immediate, catastrophic failure of the No. 2 engine. As evidenced by crew processes up to this point, these novice crews have varying difficulty with setting and maintaining goals, adjusting to requirements of the operational tempo, retrieving and applying basic systems knowledge, time management, and task management skills.

The control tower transmission concerning the truck at the intersection of the runways was in effect *firewalled* (i.e., a strategy used to intentionally delay or prevent incoming information (S. K. Hunt, personal communication, February 2, 2002)) by both the pilot and the COTAC of Crew 16. This very experienced crew had never flown together. Yet, they developed a mutually shared mental model of the desired result that prompted both crewmembers to apply an apparent "non-receive" mode to some stimuli. Crew 16 intentionally ignored the possibility that more information was available from the tower to carry out their decision plan. (This crew ranked as the second highest performers of the sixteen crews in the outcome ranking. In the debrief, the pilot indicated that he was unaware of the tower informing them of the fouled runway condition.) Although luck may have played a part in the successful outcome of this crew, the information delaying strategy played a significant role in their final outcome.

Appendix L presents the strategies used in the final scenario events by the highest (Crew 13), mid-range (Crew 11), and lowest (Crew 6) crews in the outcome rankings.

The findings from these crews, as well as the other thirteen crews, are presented below in a composite overview of crew characteristics related to decision-making strategies that had the greatest bearing on overall performance in the study scenario.

Critical Characteristics and Factors that Define Most Successful Performance Outcomes

Findings related to the third research question show there are multiple commonalties in the crew characteristics and functioning in a number of areas of performance related to the most successful crews' decision-making strategies and patterns. Among the higher performing crews, that is crews ranked in the top quadrant in both sum process scores and outcome rankings (i.e., Crews 16, 15, and 13), the following patterns were observed in the areas most directly related to overall performance in this study scenario.

Decision planning. The most successful aircrews set up a focused strategic game plan, with a firm commitment to land, immediately after their first in-flight emergency presented itself and accelerated their activities in order to expedite their return to the field. The most successful crews appeared to visualize or imagine their desired results and worked backward to design the requirements to get there. The best performers synthesized their experience and knowledge structures to meet the scenario requirements.

Information gathering and use. The best performing aircrews: (a) shared responsibility for efficiently gathering and handling selective information, prioritized incoming communication by immediate workload and its relevance to the strategic goal, quickly executed Checklist memory items (e.g., selected portions of a few checklists were reviewed silently by COTAC's or intentionally skipped), and completed remaining procedural items; (b) did not expend energy or time pursuing new, readily available, or

irrelevant information unrelated to the desired outcome; (c) resolved ambiguous information; (d) interwove task execution and communication with brief periods of cross talk to keep each other apprised of aircraft configuration changes, position relative to the field, checklist status, and upcoming actions; (e) used consensus building at each major decision point; (f) expressed high confidence and comfort level with self and crewmate; (g) did not verbalize a wave-off as part of their contingency plan(s) as they had determined early on that this was an unrealistic option.

Pace of activities. The pace set by the highest performing crews was expedient enough to "stay ahead of the jet" but not rushed. There was no delay in the decision to return to handle the existing emergency and configure the aircraft for immediate landing at the field. Multitasking was also handled efficiently with a division of labor, as was "protection" and back up for the other crewmember as necessary to render assistance and prevent task overload that might have led to slips and mistakes.

Time awareness of communications and task requirements. High performing crews were acutely aware of time elapsed (from the initial recognition of loss of the No. 2 oil pressure) and time remaining in terms of their perception of how long an engine could continue to operate in a low oil pressure condition. They mapped time available into their perception of current state of the environment (e.g., airspace, aircraft state, etc.). They devoted cognitive resources to maintaining high levels of situational awareness to maintain focus on crucial existing and impending tasks. Better performing crews were driven by this heightened cognitive state so they could allocate the appropriate amount of time to exchange information at any point in or phase of the flight. There was a cadence of strategically timed cross talk throughout the flight. Conversely, they delayed acting on

or receiving inbound information if they felt it would distract them from focusing on crucial existing or impending tasks related to the accomplishment of their clearly defined strategic goals (or game plan). High performers remembered and stored data to share during the occasional seconds between tasks or during periods of low workload. This was accomplished by anticipating opportunities to communicate or receive information the next chance they got in the flight. Only a clear mutual vision of all necessary tasks allowed this synergistic exchange.

Crew coordination/cooperation. Teamwork and leadership were also keys to successful overall performance. In the S-3 community the pilot is designated as the aircraft commander responsible for the overall safe conduct of a flight. Crew coordination reflects the notion that the crew is a team and a "we" mentality existed in the most effective crews. For example, in several instances, the pilot of Crew 16 gave positive reinforcement to the COTAC's scan and back-up actions. In general, crewmembers were responsive to the other's judgment and requests. If a crewmember's judgment was questioned or clarification was required, it was done in a professional manner.

Risk assessment and management. Experienced crews rapidly and accurately assessed both probability and severity of the risk and made the strategic decision to land as soon as possible. They continued to evaluate compounded risk as the variables changed while effectively prioritizing a hierarchy of hazards that allowed then to deal with the most severe hazards first. High performing crews used correct application of systems knowledge in risk taking (e.g., awareness that the re-start of the No.1 engine would dramatically reduce the risk of losing the aircraft). The most successful crews also

determined that landing the aircraft with a fouled deck was a better option with less inherent risk than attempting a wave-off and they did not take time to plan for contingencies in their final decision to land.

Familiarity with the landing environment and aircraft performance capabilities. The set up for approach (aircraft positioning) to land at the very beginning of the runway played a critical factor in the outcome of this scenario. High performing crews used their knowledge of the distance to the runway, runway length, and intersection locations to make strategic decisions. The most successful pilots/aircrews displayed an instinctive ability to accurately perceive trends in aircraft's airspeed, altitude and relative distance to the intended point of landing. The development of, and the reliance on, this sort of perceptual skill set is integral to precise, safe landings with any aircraft, let alone one experiencing mechanical difficulties. Pilots that flew better-controlled approaches had the ability (both in terms of motor skills and highly developed perceptual skills) to establish and maintain desired descent rates and approach speeds. This allowed them to target the end of the runway (versus the normal landing point some 700 feet beyond the runway threshold) once informed that the arresting gear was out of service.

Crews that were in a position to land and did so adapted the standard S-3 approach pattern and deviated from the primary landing aid (i.e., *Fresnel* lens) used to guide pilots in for a simulated carrier approach on the landing field. This lens consists of five lighted cells that indicate the relative glide slope position (high, on target or low), and is commonly referred to as the *ball*. The ball is used both on a carrier deck and on landing fields (for training purposes) to guide pilots to the third of four wires available to catch the aircraft's tailhook as it approaches the landing surface. Figure 9 roughly

depicts the differences in aircraft carrier and field landing aircraft arrestment points using the *Fresnel* lens. On Runway 29, the lens provides vertical guidance to a point well beyond the runway threshold.

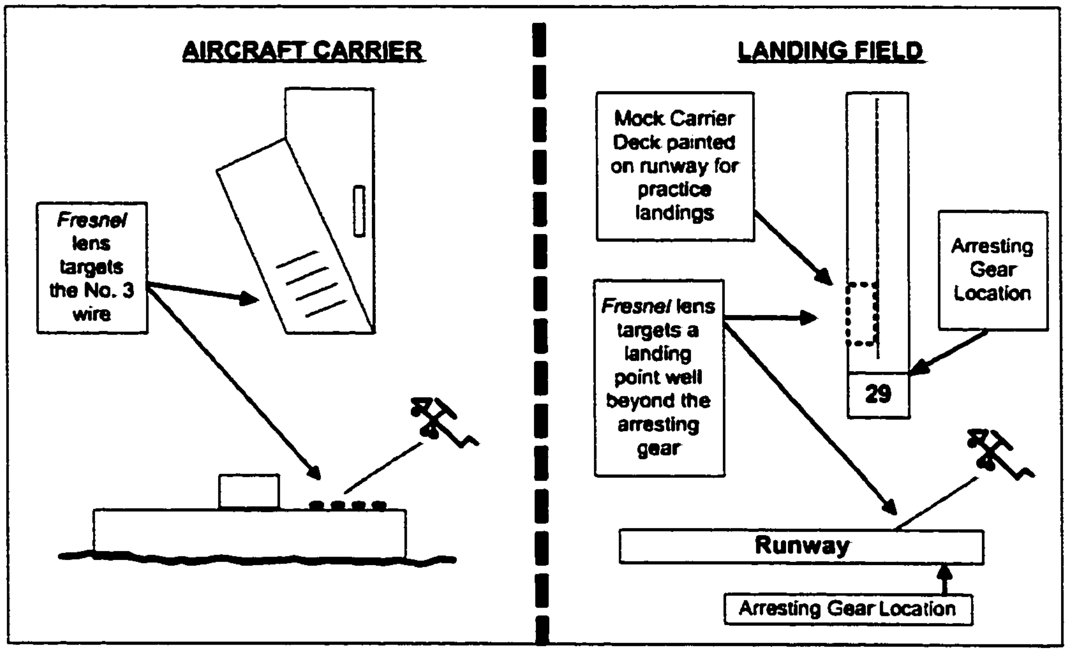

Figure 9. Comparison of target points from side and top perspectives of landing approach paths on an aircraft carrier and landing field using a *Fresnel* lens.

Generally, the more experienced pilots elected to disregard or make adjustments to fly the ball low as soon as they recognized the need to touch down with the maximum amount of pavement between the aircraft and the intersection. The better performing pilots were then able to utilize the 720 feet of otherwise "over flown" runway that was typical of an approach flown using the *ball* for vertical guidance. (S. K. Hunt, personal communication, February 19, 2000).

Summary

Data analysis revealed that individuals and teams who had the ability to use their domain knowledge and experience to recognize the need to break away from procedural and analytical process rules to cope with naturalistic contexts had better process scores and outcome ranks. Both process and outcome rankings were good discriminators between novice and experienced crews although process and outcome correlation seemed to be limited to experienced crews. The study findings clearly reflect the skillful use of advanced cognitive processes by high performing crews to adapt information gathering and decision making strategies to dynamic situations. The quantitative analysis found statistically significant differences in process performance in four scenario process elements between the two levels of flight experience represented in the groups. Based on this finding the qualitative inquiry focused on identifying and understanding the distinctions that characterized the varied performance levels.

The qualitative investigation found that there were indeed clearly different approaches to decision- making strategies including risk perception and management in better performing crews. Most experienced crews described adaptive strategies they used to rapidly identify and prioritize relevant risk factors that required immediate response. On the other hand, most novice crews continued to use procedural and analytical decision-making strategies under real-time, dynamic situational demands and overlooked the cognitive adjustments required to carry out their initial plan to land. Ultimately, the crews' perception of risk was predicated on their perception of the circumstances.

CHAPTER 5

SUMMARY, DISCUSSION, IMPLICATIONS FOR TRAINING AND RECOMMENDATIONS FOR FURTHER RESEARCH

In addition to a summary of the study, this chapter synthesizes the cross-case findings from the analysis of the main study populations and discusses the consistency of these findings with the theoretical frameworks of naturalistic decision making researchers. The chapter concludes with the researcher's interpretation of the study and the potentially profound implications for aviation training advancements and recommendations for further research.

Summary of the Study

This empirical study examined five interrelated elements: (a) to investigate the relationship of experience on aircrew process ratings and decision results (outcome); (b) to determine any relationship between process ratings and outcome rankings; (c) to distinguish the decision-making strategies of aircrews; (d) to determine if decision-making strategy patterns of successful and less than successful aircrews support or refute theoretical concepts/models for naturalistic behavioral analysis; and (e) to identify crew performance characteristics of the most successful crews.

Summary of the Methodology

The case study approach was selected to provide meaningful data to identify the thought processes used by more and less effective aircrews. Case study comparisons sharply defined how individual and aircrews management of uncertainty differed under

varying degrees of event structure and time constraints. Data for this study consisted of instructor evaluations of 8 novice and 8 experienced aircrews consisting of one pilot and one COTAC (co-pilot tactical coordinators). Aircrew performance in a flight simulator scenario provided a realistic environment to observe and rate aircrew performance over seven events of increasing complexity and uncertainty. Digital files of the flight event were then used to cue retrospective verbal reports of crewmembers' metacognitive processes related to situational awareness and decision making. Crew outcome rankings were generated from independent rater judgments of optimum crew and aircraft disposition at the end of the scenario. The research into the underlying issues related to differences in aircrew processes and outcome was guided by results of the inferential and descriptive statistical analyses of process score and outcome ranking data.

Summary of Key Findings

Study findings, from both the descriptive and inferential quantitative and qualitative cross-case analyses, provide multiple lines of evidence towards the same conclusion and are summarized below. These findings confirm the findings of other researchers in naturalistic decision making and are interpreted as they relate to the study hypotheses and research questions in the sections that follow.

- The overall superiority of process scores received by experienced crews was statistically significant.
- Scenario events with sequentially increased uncertainty and limited response time served as good discriminators of process ratings with statistically significant differences between the groups in the last four more complex events.
- The superior outcome rankings of experienced crews were statistically significant.

- Strong indicators of a positive relationship between process and outcome ($r = .64$) was limited to the experienced group. There was no strong process/outcome relationship found with novices ($r = .16$).

- Debriefing/interview data provided specific instances of thought protocols that revealed differences in strategy selection and application between higher and low performing aircrews.

- Higher performing crews demonstrated better ability to use adaptive strategies to identify relevant concerns, to evaluate risk, and to develop a practical solution with no increase in effort within the time available.

- Less experienced or poorer performing crews were driven by procedural and analytical concerns at inappropriate times and did not make the cognitive adjustments required to relinquish a linear systematic approach for flying when in an extremis situation.

Hypothesis 1: Experienced crews receive better process score ratings than novices in handling a specific in-flight emergency involving uncertainty

Process ratings. Quantitative analysis centered on a two-factor (experiential) multivariate analysis of variance, with seven dependent measures (decision points). In general, experienced crews demonstrated more consistent performance with less deviation in mean and variance differences in process scores across all events compared to less experienced crews. The last two event conditions (i.e., *no arresting gear* and *fouled deck*) clearly illustrated the importance of an aircrews' ability to recognize the need to modify their routine situational assessments and decision-making processes. There were distinct differences between performance groups in dimensions such as pattern matching, memory for domain-relevant facts, conflict resolution, risk assessment, cockpit resource management, and decision strategy and execution. The most experienced crews exhibit "clusters of skills that tend to make their performance more

stable, less error prone, and more efficient than novices or intermediates (Seamster, Redding & Kaempf, 1997, p.29).

For example, the high performing crews executed a strategic game plan without major re-analysis of their plan every time the scenario context changed and engine time remaining became more uncertain (i.e., low oil pressure in remaining engine, arresting gear not available, and truck in runway intersection). In general, experienced crews were able to draw on their experience and domain knowledge to intuitively recognize the need for creative adaptive action and adjusted their flight paths to land without use of arresting gear. Seven of the eight experienced crews deviated from procedural requirements for an arrested landing and touched down at the closest possible point on the runway. This was a risk tradeoff between eliminating "room for error" provided by a normal approach path and giving the crews the use of additional runway to stop prior to the truck in the intersection (i.e., fouled deck).

Experienced crews with hundreds of carrier deck and airfield landings possess the confidence and knowledge that with adjustments to the landing approach, the aircraft was capable of stopping prior to the intersection. Conversely, inexperienced crews tended to fly less disciplined (i.e. less controlled) approaches with higher unintentional deviations in airspeed, altitude and approach paths (not to be confused with glide paths) than their experienced counter parts.

Seven out of eight novice crews that had established a game plan to land elected to wave-off. When the arresting gear was unavailable most novice crews could not "break set" with the routine field arrestment landings they are required to execute in all training events. Unlike the more experienced crews, most novice crews did not take into

consideration the uncertainty of the circumstances in their risk assessment and management of the situation. These crews subsequently planned their approach path without considering there may not be an opportunity for a second chance to land. Four out of five novice pilots did not follow through with their intended plan to land, even without the arresting gear, because they realized they had not flown a precisely controlled approach pattern (i.e., did not have the automated basic domain skills (i.e., stick and rudder skills) to make the last minute correction). Furthermore, the inexperienced crews had problems retrieving and using information under stress about runway available until the intersection, and/or lacked the confidence to follow through with their intentions (i.e., several of the pilots took last second direction from the COTAC to wave-off).

Hypothesis 2: Decision results (outcomes) of experienced aircrews are rated as higher quality than those made by novice aircrews

Outcome rankings. The complexity of the decision problem in the final decision to land or eject was the major determinant of outcome ratings. Inexperience with a compound emergency in an unstable landing environment and inexperience in dealing with time-critical situations prevented most novice crews from making the optimum decision in this case. The interpretation of the experiential factors and patterns that relate to the required cognitive effort and time involved in this decision process is summarized as follows: (a) experienced crews did not consider options to deviate from their original strategic plan to land immediately, and (b) inexperienced crews more often proposed and considered more than one option when landing gear was de-rigged and then again when an obstacle was placed in the runway.

Hypothesis 3: Aircrew process ratings are correlated with the decision outcome ratings

Process scores' inter-relationship with outcome rankings. Analysis of process and outcome scores showed a main statistical effect for flight experience. The study results suggest that experience was a significant factor in aircrews that received higher scores in the scenario's later decision points. The experienced crews generally achieved higher process score totals as well as higher outcome scores from independent rankers. No strong apparent relationships were found between novice process and outcome scores. However, there was a positive correlation with experienced crews between outcome ranking and two process events (no *arresting gear game plan* and *foul deck game plan*). Process and outcome rating analysis revealed several important aspects related to overall performance. The major findings from deconstructing performance attributes of high and low performing aircrews that support quantitative results in terms of ability to predict process and outcome performance are: (a) the most notable distinction in aircrew performance occurred in less structured decisions under severe time pressure that required an adaptive response to the decision problem to satisfy immediate safety of flight concerns, and (b) there exists a longitudinal effect of an aircrew's strategic, tactical, and procedural planning and execution, beginning in the brief and continuing throughout the flight.

The researcher used empirical generalizations made from quantitative analysis beyond the significant effect size itself to guide the qualitative inquiry. Since the difference in process scores between experienced and novice crews were expected, the research questions focused on the reasons for these differences. Although rule-based decisions were applied well by almost all crews throughout the scenario, the performance

of most novice crews deteriorated when unanticipated problems demanded creative thinking during the latter stages of the scenario.

The evidence clearly indicates that experience is a primary factor in the ability to: (a) concisely recognize and assess risk in atypical situations; (b) apply domain knowledge, skill and experiences to perceive time available to save the aircraft and crew; (c) adjust a plan to respond to situational dynamics, and (d) consider only one option at a time. These findings are supported by decision research findings: increases in task complexity drive altered evaluation and choice strategies used to make both high and low risk choices (Payne, 1985).

Research Question 1: What adaptive recognitional/metacognitive decision-making strategies emerge from aircrews?

Emergent recognitional/metacognitive strategies. Beyond domain knowledge, expert performers use meta-recognitional skills that include rapid information search and prediction of option success to inform adaptive decision-making processes. Key research findings throughout naturalistic decision research summarized by Means, Salas, Crandall and Jacobs (1993) support the description of the strategies reported by the aircrews to simplify complex problems while maintaining flexibility in thought processes as well as reactions to unfolding events. Adaptive recognitional/metacognitive strategies were used to systematically search for information, and quickly identify, characterize and frame problems.

Crew use of adaptive strategies supports naturalistic decision making theory that the most successful crews would have employed adaptive strategies as time pressure and

need for more immediate action grew stronger as the scenario progressed. Interview analysis revealed that adaptive "satisficing" or "good enough" (Simon, 1955) strategies emerged in the more uncertain and time critical portions of the study scenario to effectively achieve immediate short-term and long-term goals. These decision-making strategies characterized by the effort and accuracy trade-offs in the more progressively uncertain decision-making environments in the study scenario were consistent with the findings of Payne, Bettman, and Johnson (1988, 1990).

Data analysis yielded two multi-dimensional matrices to synthesize the decision strategies employed by study aircrews under different levels of uncertainty. The conceptual categories used to develop the matrices were based on the ideas of Klein, 1983; Payne, 1985; Humphreys & Berkeley, 1985; Endsley, 1995, 1997; Orasanu et al., 1993; and Cohen, et al., 1996. These matrices proved useful in identifying critical links in the crews' performance in the study scenario during increasing time constraints and uncertainty. Variations in the range of adaptive strategies used by individuals/crews for information gathering, processing and decision making are included in Tables M_1 and M_2 of Appendix M.

Figure 10 graphically portrays the range of responses and crew interactions/ actions related to the cognitive efforts of aircrews in a multi-crewed aircraft in the study scenario. Process data used to develop this explanatory model evolved from the data compiled for building the matrices in Appendix M. Data summarized included a priori crew explanations of thought processes and actions regarding processes of information gathering, goal development, risk assessment, and subsequent technical, tactical and strategic decision making under various conditions of uncertainty and time pressure. The

explanatory model captures the decision-making temporal considerations and associated level of interaction between crewmembers and with external sources associated with updating situational awareness, risk assessment, strategy revision and implementation. Various strategies depicted in Figure 10 were used by aircrews to manage uncertainty and are identified throughout the decision cycle starting in the top left corner with the label *new information*.

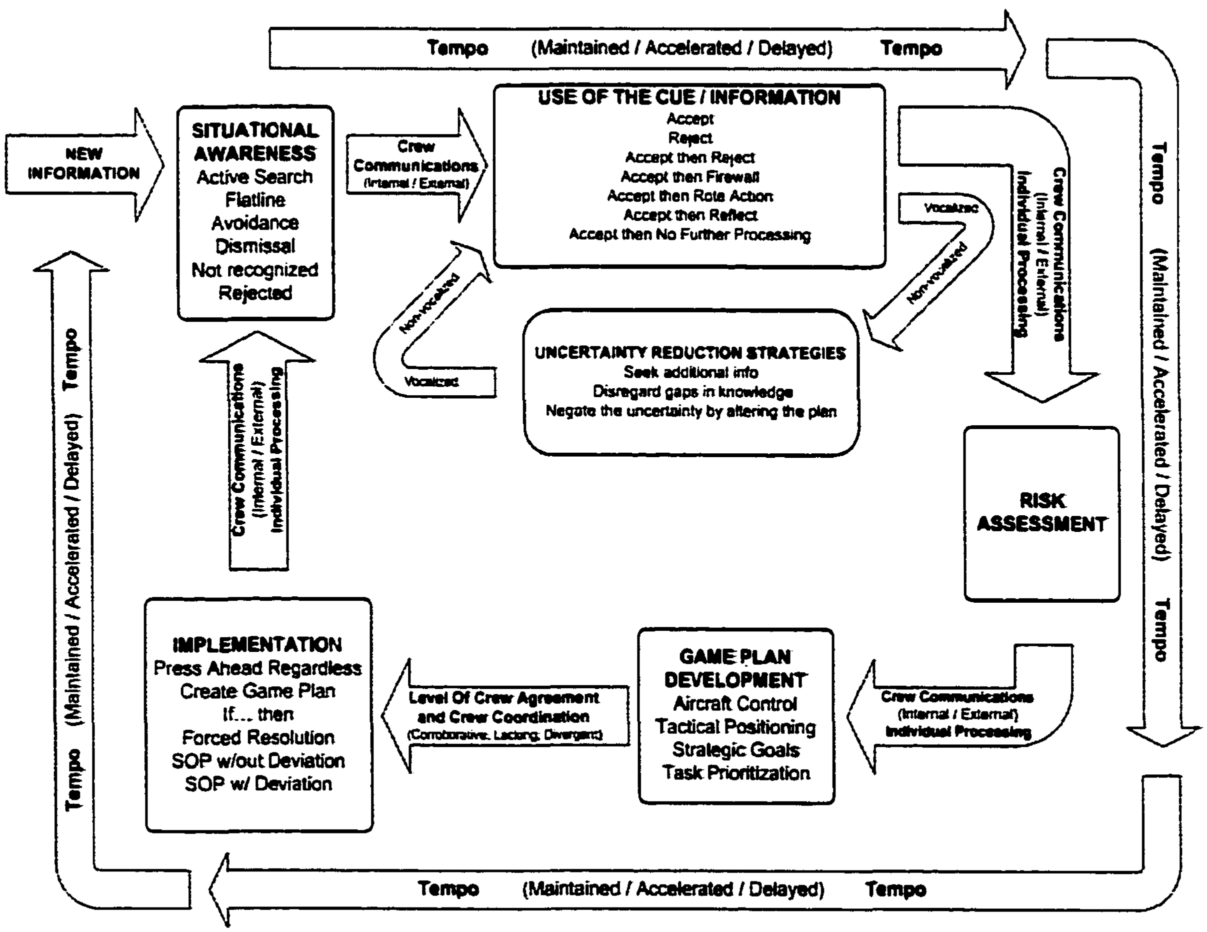

Figure 10. Explanatory model of the range of responses associated with aircrew decision making in a scenario involving naturalistic conditions

The findings support that the effective use of critical thinking skills and processes as framed in the Recognition/Metacognition (R/M) model (Cohen, Adelman, Tolcott, Bresnick, & Marvin, 1993; Cohen, Freeman, & Wolf, 1994) resulted in superior outcomes. The explanatory model created with study data, illustrated in Figure 10, supports Cohen's focus on "an integrated picture of how knowledge structures are created and adjusted in dynamic environments" (Cohen, 1993a, p. 49). The findings of this study support that success of the processes used to manage uncertainty were dependent on the complexity of the problem "manipulated through variations in the number of alternatives in the choice set, the number of dimensions of information (attributes or outcomes) used to define an alternative, and the amount of time available for making the decision" (Payne, 1985, p. 7). The next section will discuss the way in which the crew processes identified for managing uncertainty influenced performance in the scenario.

Research Question 2 : In what way do adaptive/metacognitive decision-making patterns differ among successful and less than successful aircrews?

Differences in decision-making patterns of high and low performing crews.

Interview analysis revealed that recognitional/metacognitive strategies were required in the more uncertain and time critical portions of the study scenario. Aircrews that could quickly identify the first option in a sequence of options that would immediately "satisfy" requirements to make a safe landing had better outcomes. Aircrews that had problems prioritizing and/or tried to justify their decision-making process (in several cases to authorities rather than themselves) by falling back on standard operating procedures and/or weighing multiple options (often using biased probabilities or logic) were not as successful in optimizing their outcomes. For certain tasks in the scenario, as well as in

operational settings, procedural and analytical strategies will result in better performance (Klein, 1997a) but this scenario was designed to force the use of adaptive decision making for the best possible results.

Experienced aircrews were more often and better able to use past experience to adapt interactions between crewmembers that resulted in quicker situational assessments and better decision-making focus to achieve optimal results. As depicted in the Recognition-Primed Decision Model (Klein et al., 1986) and used by experts in various field studies (Klein, 1997b), this adaptive critical thinking process focuses on the sequential evaluation of options for immediate "best fit" in a dynamic situation rather than a time-consuming weighing of options for the best possible solution and/or preparation to justify actions.

With regard to satisfying the immediate landing requirements in the scenario, the core differences in performance between novice and experienced aircrew outcomes included: (a) more experienced crews were better calibrated in the strategic outcome goal and the closer they got to the threshold they focused narrowly and sharply on the commitment to the landing option alone (i.e., only information pertinent or germane to a accomplishing a safe landing was processed; (b) novice crews tended to discuss/investigate non-landing options and even invented additional non-landing options while on the final approach to land. Novice crews did not select or correctly apply appropriate decision strategies for this final event involving a truck on the runway during final approach.

The rule-based decision strategies used by most novice crews were not flexible enough to work in a novel situation. The procedural analytical, option-weighing

strategies were too time consuming and/or inappropriate probabilities of error/success were placed on options. Interview analysis revealed that recognitional/metacognitive strategies were employed by the more experienced crews to guide, limit and stop information search in the final time-critical events that called for satisficing. This finding supports the use of critical thinking strategies for quick retrieval of assumptions and identification of relevant information; that have been identified as: (1) critiquing or accurately evaluating/characterizing the problem, (2) monitoring a course of action to assess whether the methods and results of the decision process will be satisfactory, and then (3) correcting or regulating the plan with a sequential evaluation of options with a commitment to the first acceptable alternative rather than trying to optimize by waiting for analytical results (Klein, 1993; Rouse & Valusek, 1993; Cohen et al., 1996).

Since a "vital element in all strategies is a specification of both the amount and order in which information is processed" (Maule, 1985, p.71) the information gathering and use by aircrews was of primary interest in the data analysis. The finding that the individual/crews' perception of their circumstance had a major effect on their risk perception of a situation led to the investigation of information processing strategies and judgments. Errors in information gathering strategy or errors in use of strategy to evaluate and characterize the situation or problem were a major determinant of the final outcome in the scenario performance. Aircrew uses of judgmental rules, known as heuristics, were used to break down difficult tasks into simpler ones. Although the use of heuristics is "valid in some circumstances, in others they lead to large and persistent biases with serious implications for decision making" (Slovic, Fischhoff, & Lichtenstein, 1982).

Lower performing crews were poor at information gathering to use for strategic judgments. They commonly did not work at "staying ahead of the jet" (i.e., planning and identifying potential risks in light of their strategic goals). Novice crews routinely became involved in completing the tasks at hand (e.g., orbiting in the vicinity of field to complete troubleshooting drills/checklists) and attempted to make even more time available (e.g., attempted to execute the full instrument approach procedure rather than taking vectors or commencing a visual approach) for these relatively irrelevant tasks. Novice crews also devoted cognitive capacity to inappropriate concerns and demonstrated an over-reliance on and unfamiliarity with checklists. One of the most obvious differences between novice and experienced crews in this scenario seems to hinge on their willingness, or lack thereof, to abbreviate, deviate from, or in some cases completely ignore checklists items. Once the experienced crews ascertained the need to land immediately there was a relentless concentration on getting the jet "on deck." Novice crews, on the other hand, never really seemed to recognize or generate the same sense of urgency "to get at least the important stuff done" that was repeatedly seen in the experienced crews.

Although most novices displayed many attributes of good performing crews they were either not consistent and/or they were driven by procedural and analytical concerns at inappropriate times. Applying rule-based solutions that they thought would help to define and control the situation drove many novice crews. There were multiple examples of novice crews substituting dogmatic compliance of checklist for the intent of the checklist (i.e., flight safety). For most novice crews, the regard for prolonging the flight with "clean-up" issues outweighed the urgency to get the aircraft on the ground. For

Crew 1, this mindset became extreme to the point of substituting checklist compliance for aircraft control. In the debrief the COTAC explained the crew's priorities for "taking the time to go through all the checklists" and remarked, "...we just need to get through the landing checklists and get them done *otherwise* this plane is not going to fly anymore."

As the number of competing, high priority tasks began to increase in both size and importance experienced crews very clearly demonstrated a "triage-like" task management mentality. There was a need to address everything at a *satisificing* level in the time available. Accordingly, these crews relied upon their experience and well-developed sense of judgment to cut corners wherever and whenever appropriate as they performed safety of flight tasks such as checklists, systems monitoring, troubleshooting, navigation, and coordinating with the control tower. Experienced Naval aviators recognized this situation required a "gear, flaps, hook, land" mentality in which all other checklist items become secondary. Their communications with the control tower became somewhat terse directives (e.g., "We'll be taking a trap on 29" vs. the normal request for landing advisories or clearances) and the general concern for their equipment became much more "survivalist" vs. "maintenance friendly" in nature. The willingness of experienced crews to restart, and quite likely completely destroy a two million dollar engine in order to save a $30 million aircraft is a perfect example of this type of "triage" task management. Everything that was important received attention, obviously not as much attention as was ideal, but at least enough to get the jet landed on the runway.

The process scores reflected a degree of aircrew accuracy in identifying relevant cues, organizing the information into a judgment and then using appropriate decision-making strategies over a series of judgments. According to Arkes and Hammond's

model of judgment analysis, "judgment is a cognitive process similar to inductive inference" (1992, p. 7). Since high judgmental accuracy is considered an essential attribute of high performance it was expected that low judgmental accuracy would be reflected in lower performance and scores. Sample quotes representing a crewmember's use of heuristics, developed as a result of various studies in judgments of probability, are presented in Table 11.

Table 11

Sample Heuristics used in Risk Assessment by Aircrews

Heuristic (rule)	Crew /position	Example of Use in Predicting Risk
Imaginability- novel situation evaluated by imagining contingencies (with no experience to use) (Tversky & Kahneman, 1973)	Crew 2 Pilot Waved-off 9/16 in outcome ranking	" We didn't have short field gear so that's a factor. We might have been able to stop; we might not. So they weren't going to get the truck moved so we would probably have hit it if we landed. There's no way I could have stopped it there."
Availability- ease of retrieving instances/occurrences to assess probability (Tversky & Kahneman, 1973)	Crew12 Pilot Waved-off 15 of 16 in outcome ranking	"I don't think I would have stopped in time. I think I might have been able to stop it. However, that's not always the case even when I slam on the brakes. That was 50/50. I would say from my experience landing in this jet that the brakes aren't what they ought to be. That would have been a 50/50 chance of us plowing into the truck. So weight the odds of that if the other engine can keep going and both those chances are real bad."
Illusory Correlation- over-estimate of strength of associative bond (Chapman & Chapman, 1969)	Crew 6 Pilot Waved-off 16/16 in outcome ranking	"Well, I wasn't thinking about how far the intersection was. All I heard was 'intersection' in my mind. Now sitting here I can stop and think about how many times I've stopped the jet before the intersection. But I was thinking along the lines of conservatism, I guess, and not sure whether it could stop in time."
Evaluation of Conjunctive- likelihood plan will succeed through series of events (Bar-Hillel, 1973)	Crew 11 Pilot Landed 8/16 in outcome ranking	"You know if we can't stop by the time we get to the intersection then most likely we're not going to be able to stop. There's the long field gear but there's some other kind of problem [involved]. So, this jet easily stops within 4,000 feet on runways. If we weren't stopped by that point you're slow enough to go off or at that point you're ejected."

In conclusion, the major issues related to poor performance included: (a) unresolved issues in learning/training (systems knowledge, technical skills, crew coordination, etc.), (b) inability to improvise when safety of flight called for abbreviated or expeditious handling of checklists, (c) fear of not being able to justify a decision to deviate from published procedures or regulations when warranted by airborne emergencies, (d) inability to strategize to keep the goal with the biggest payoff in focus, (e) lack of confidence in their technical skills, and (f) lack of assertiveness as pilot in command.

Crew characteristics and factors that defined the most successful crew outcomes. The attempt to develop a stereotypical definition of an effective crew by studying the range of performance between novice and experienced crews was more complex than originally anticipated. A consolidated list of attributes related to different performance levels was revealed through various data sources in the study. Attributes of high performance were found in lower performing crews but not to the extent and consistency found in better performing crews. Poor performance characteristics were also found across all crews. In general, novice crews exhibited more of these characteristics than experienced crews as evidenced by the process and outcome rankings. A summary of specific crew characteristics related to performance levels is located in Appendix N.

Substantial flight time alone was not a qualifying factor in distinguishing "expert" performers in the study scenario. Two cases underscored that other factors may be involved with flight performance beyond, or in spite of, an accumulation of flight hours. The crew with the least amount of total flight time in the S-3 exhibited many of the characteristics of expertise and team skills associated with aircrews at the other end of the

flight experience spectrum. The essential factor in successful performance by this novice crew was the ability of each crewmember to adjust his cognitive style as well as his workload prioritization to meet emerging requirements. For work in a crew environment, this meant that as context and task characteristics changed each crewmember relinquished/adjusted his viewpoint or workload as necessary. Traditional teamwork includes balancing positional duties in a "divide and conquer" approach with "back-up" for procedural compliance and situational awareness (i.e., aircraft altitude, airspeed, and positioning relative to an acceptable approach path profile). In the case of the poor performing experienced crew, the pilot's overbearing attitude towards the COTAC whom he outranked and complacency in verifying identification of problems or procedures (e.g., started to shut down wrong engine, never associated smoke with low oil pressure, elected to wave-off) were not representative of an experienced fleet aviator.

The higher performing crews displayed more of these traits and were more consistent in demonstrating these expert capabilities than the average or lower performing crews. The goal of the cross-case study analysis was to find the extremes of performance and create a stereotypical definition of a good crew. Therefore, case studies of aircrews ranked highest, mid-point and lowest on the process and outcome score continuum were analyzed to reveal differences in areas found to be primary determinants of process and outcomes. Performance differences that grew progressively greater as the difficulty level of the scenario increased were found in the following areas: (a) knowledge and skills involving aircraft control, (b) leadership, (c) cue recognition, identification and response, and (d) strategic focus aligned with tactical and procedural focus during each phase of flight or during an abnormal event with competing issues.

The experienced crews reacted differently to apparent inconsistencies in the situation and did not deviate from their original strategic plan as the novice crews did on the final decision. As reflected in the Recognition-Primed Decision (RPD) model, a comparison of options was typically not done by experienced personnel in the final scenario events. Once they entered the "end stage' they remained with their original plan to return to base and land as soon as possible regardless of the distractions presented to them in the scenario. Inexperienced crews more often created options when informed that the landing gear was de-rigged and then again when informed that there was an obstacle on the runway. This supports the notion that novice crews' prior training and inexperience in landing at the field under extremis situations was a major factor in their inability to make a choice from dynamic sequential options. As one novice pilot explained, his reliance on procedural responses remains consistent in the transition from a training command single-seat jet, where "pretty much if things don't work perfectly- eject" to piloting a multi-seat jet where "you definitely change how you handle emergencies . . . to improve crew coordination, taking more time with your procedures to make sure you get them right, and not place yourself in more extremis can make a big difference."

The study findings support that flight experience is the major determinant in a crew's ability to react more quickly and more accurately to complex situations involving uncertainty and severe time constraints. This conclusion is supported by previous research findings using gambling predictions (Payne, 1985; Payne, Bettman, and Johnson, 1988) that a decision-maker has a multitude of strategies to select from to predict outcomes depending on the trade-off between costs and accuracy given

constraints of the situation. The metacognitive processes shared by the crews during the debrief/interview suggest that a crewmember's personal sense of confidence as well as his comfort level with the other crewmember's ability to back-up and carry out the specific tasking requested is important to effective team performance. This required crewmembers to deal with some "distracting" concerns by internalizing concerns, alternate options, etc. Of equal importance was communicating succinctly and clearly at appropriate times. In one case, an experienced COTAC internalized his processing of the probabilities of missing different runways at the field. He then continued to use metacognitive processes to plan contingencies as a means to allow the pilot to focus on making final adjustments to land the aircraft. With poorer performing crews, there were many instances of a crewmember relaying concerns in inappropriate ways and times that negatively affected strategy processes and resultant outcomes.

Implications for Training

This study context reflected a realistic context to investigate the use of adaptive/recognitional strategies that approximate the accuracy of normative rules with substantial savings in effort. Study findings supported that the use of well-defined behavioral and cognitive constructs provide a more robust approach to aircrew evaluation and training feedback than skill-based training guided by post-hoc analysis of mishap data. More defined decision problem representations allow for identification of the aircrew attributes that really matter in performance at both the novice and expert levels and capture conditions conducive to human error.

Cognitive dimensions added to observable evaluation criteria created both a multi-dimensional evaluation and a crew self-assessment tool. Technical skills were

complemented by metacognitive attributes such as goal identification, ability to assess effort and accuracy required of various strategies, and consequences of actions or options not selected. Fleet crews exposed to the challenging study scenario as a part of their NATOPS qualification all agreed that the scenario, in concert with the more in-depth debrief, provided a better learning experience than provided by routine qualification evaluations. With this multi-faceted approach to evaluation, training objectives can be improved to reflect better-defined essential behaviors and processes for both individuals and teams. Improve and standardize events and performance evaluations focused on critical thinking skills will promote solution-oriented interactive briefs, focusing on specific behaviors keyed to training goals.

Metacognitive and meta-recognitional focused verbal protocols are a great addition to the debrief as well as the classroom discussion and "provide accurate record of how an individual internally represents ideas, and in certain situations provide an appropriate measure of information processing" (Simon, 1979, p.69). Use of cognitive oriented questions by the instructor involves asking the crewmembers about the potential as well as the actual impact of their thoughts and actions so they can generalize lessons learned beyond a particular scenario. The debriefing protocol, including the study cognitive probes, is now routinely used by instructors at the S-3 Fleet Replacement Squadron (i.e., training command). Instructors have found that these types of questions provide both them and the aircrew with a better understanding of the crewmember(s) underlying processing activity.

Scenario-based training with cross-case analysis is also useful, as in this case, to identify possible improvements in operational performance. Additionally, decision trees

built from crew performance data can be a useful format to ensure correct and rapid execution of immediate action and non-normal procedures in flight manuals and checklists.

Recommendations for Future Research

There has been considerable research in the area of multi-crew/team training yet cognitive approaches to team training have received very little empirical attention. The adaptation of available empirical data for practical use and in training warrants further study. For example, in this study the variables in the data collection instrument were sufficient to measure the constructs of the study but findings reveal that there are more defined constructs available to capture team requirements for decision-making processes in naturalistic contexts. Although traditional training data collection and evaluation practices identify the behaviors and outcomes associated with performance issues they do not commonly seek out efficacy and accuracy of the underlying thought processes.

The generation of more complete data on attitudes, cognitive processes, and skills in a realistic, challenging context will result in identification of data categories that are more significant to aircrew performance assessment and feedback for all levels of training and operations. Further research using naturalistic decision making models and theories incorporated in the research design, using both quantitative and qualitative comparisons of domain experts/novices in realistic real-time events, will add relevant details and issues related to the process of decision making for both individuals and aircrews. To date, cognitive task analysis to study expert/novice differences has been generally limited to elicitation of past events with experts in a domain (Hoffman, Crandall, & Shadbolt, 1998).

The use of a realistic context to study decision making in the field provides opportunities to elicit memories immediately after an event rather than days, months or years after an experience. For example, the inclusion of novices in a task analysis can generate data to better define problems to be addressed in training. Probing questions that engage the novice to think about his inferences, motivations, attitudes, and coordination responsibilities related to the decision-making process highlighted areas not routinely addressed in the curriculum (e.g., different mental representations and assumptions of the problem that led to opposing problem solving strategies, overt or covert deliberate disregard for the other crewmember's rationale for a plan of action, inability to behave adaptively to trade-off standard requirements for an effective level of effort to satisfy immediate operational requirements, etc.).

Employing a digital simulator data collection, debrief and data analysis device to capture events by time or category is an invaluable aviation-research tool that combines "real-time" aircrew and aircraft performance data. Immediate retrospective recall by participants provides more opportunities for participants to verify the timing and context of events. The timeliness and availability of details of this approach to data collection provide more opportunities for an aviator's robust reflections on decision thought protocol and/or insight into a particular cue, judgment, use of analogues, plans, options, etc. as well as unrecognized potentially dangerous precursors to those events (Klein et al., 1986).

Summary

This chapter provided a final synthesis and summary of findings related to the hypotheses and related research questions and discussed the consistency of these findings

with the theoretical frameworks of naturalistic decision making researchers. The focus of this type of decision-making analysis follows on Cohen's (1993b) argument that formal decision making models do not capture the adaptive characteristics of real-world behavior and that "improvements in decision making need not require imposing analytical methods" (p. 99). Although lack of comparable studies prohibits precise comparisons, this interpretation of both the statistical and qualitative data supports the data and theories of other investigators studying novice and expert performance in aviation and other domains. Naval aviation has proved a useful area for studying the relationship between experience and how an individual selects a strategy based on the problem context. The study of aircrew performance process and outcome differences in terms of their operational decision-making abilities provided insight into the essential structure of various motivations, acts, choices, and decisions made by aircrews that both reflect and contribute to decision-making theories. Multiple methods of analysis provide better prospects for greater understanding of aircrew decision making, judgment, and problem solving skills (e.g., considerations of feasibility, constraints and relevant tradeoffs) and is a distinctive approach to gain insight and understanding of adaptive approaches to decision making from multiple perspectives.

REFERENCES

Arkes, H. R., & Hammond, K. R. (1992). Judgment and decision making: An interdisciplinary reader. Cambridge: University of Cambridge.

Baker, D. P., & Salas, E. (1996). Analyzing team performance: In the eye of the beholder? *Military Psychology*, 8, 235-245.

Bainbridge, L. (1999). Processes underlying human performance. In D. Garland, J. Wise & V. D. Hopkin (Eds.) *Handbook of aviation human factors*. Mahwah, NJ: LEA.

Bar-Hillel, M. (1973). On the subjective probability of compound events. *Organizational Behavior and Human Performance*, 1973, 9, 396-406.

Beach, L. R., Chi, M, Klein, G., Smith, P., & Vicente, K. (1997). Naturalistic decision making and related research lines. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic decision making* (pp. 29-35). Mahwah, NJ: LEA.

Berenson, M. L., & Levine, D. M. (1998). Business statistics. Upper Saddle River, NJ: Prentice Hall.

Bowers, C. A., Jentsch, F, & Salas, E. (2000). Establishing aircrew competencies: A comprehensive approach for identifying CRM training needs. In H. O'Neil and D. Andrews (Eds.), *Aircrew training and assessment* (pp. 67-83), Mahwah, NJ: LEA.

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

Cannon-Bowers, J. A., Tannenbaum, S. I., Salas, E., & Volpe, C. E. (1993). Defining competencies and establishing team training requirements. In R.A. Guzzo, E. Salas and Associates, (Eds.), *Team effectiveness and decision making in organizations* (pp. 333-334). San Francisco: Jossey-Bass.

Chapman, L. J. & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology*, 74, 271-280.

Cohen, M. S. (1993a). Three paradigms for viewing decision biases. In G. Klein, J. Orasanu, R. Calderwood, and C. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 36-50). Norwood, NJ: Ablex.

Cohen, M. S. (1993b). The naturalistic basis of decision biases. In G. Klein, J. Orasanu, R. Calderwood, and C. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 51-99). Norwood, NJ: Ablex.

Cohen, M. S., Adelman, L., Tolcott, M. A., Bresnick, T. A., & Marvin, F. F. (1993). *A cognitive framework for battlefield commanders' situation assessment* (Tech. Rep. No. 93-1) Arlington, VA: Cognitive Technologies, Inc.

Cohen, M. S., Freeman, J. T., & Wolf, S. (1996). Metarecognition in time-stressed decision making: Recognizing, critiquing, and correcting. *Human Factors,* 38 (2), 206-219. Santa Monica, CA: Human Factors and Ergonomics Society.

Cohen, M. S. & Freeman J. T. (1996). Thinking naturally about uncertainty. In *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting,* 1, 179-183, Santa Monica: Human Factors and Ergonomics Society.

Cohen, M. S., Freeman, J. T. & Thompson, B. B. (1997). Training the naturalistic decision maker. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic decision making* (pp. 257-268). Mahwah, NJ: Lawrence Erlbaum Associates.

Cook, T. D. & Campbell, D. T. (1979). *Quasi-Experimentation: Design & analysis issues for field settings.* Boston, MA: Houghton Mifflin.

Cronbach, L. J., (1990) *Essentials of psychological testing* (5th Ed.). New York: Harper & Row.

Ebel, R. I. (1951). Estimation of the reliability of ratings. *Psychometrika,* 161, 407-424.

Einhorn, H. J. (1980). Learning from experience and suboptimal rules in decision making. In *Cognitive processes in choice and decision behavior.* Hillsdale, NJ: Erlbaum.

Endsley, M. R. (1995). Measurement of situational awareness in dynamic systems. *Human Factors*, 37 (1) 65-84.

Endsley, M. R. (1997) The role of situational awareness in naturalistic decision making. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic decision making* (pp. 269-283). Mahwah, NJ: Lawrence Erlbaum Associates.

Flin, R. (1998, May). *Applying NDN to Critical Incident Management.* Paper presented at the Fourth Conference on Naturalistic Decision Making, Warrenton, VI.

Gigerenzer, G. & Todd, P. M. (1999). Fast and frugal heuristics. In G. Gigerenzer, P. M. Todd, & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp.3-34). New York: Oxford University Press.

Glaser, R., and Chi, M.T.H. (1988). 'Overview'. In M.T.H. Chi, R. Glaser, and M. J. Farr (Eds.), *The nature of expertise (*pp xv-xxxvi). Hillsdale, NJ: LEA

Hammond, K. R., Hamm, R.A., Grassia, J., & Pearson T. (1997). Direct comparison of the efficacy of intuitive and analytical cognition in expert judgment. In W. M. Goldstein, & R. M. Hogarth (Eds.), *Research on judgment and decision making; currents, connections, and controversies* (pp. 144-180). Cambridge, UK: Cambridge University Press.

Hartel, C. E. J., Smith, K., & Prince, C. (1991). Defining aircrew coordination: Searching mishaps for meaning. Paper presented at the Sixth International Symposium on Aviation Psychology, Columbus, OH.

Hinkle, D. E., Wiersman, W., & Jurs, S.G. (1994). *Applied statistics for the behavioral sciences.* (3rd ed.). Boston: Houghton Mifflin.

Hoffman, R. R., Crandall, B. & Shadbolt, N. (1988). Use of the critical decision method to elicit expert knowledge: A case study of the methodology of cognitive task analysis. *Journal of the Human Factors and Ergonomics Society*, Volume 40, 2, June 1998.

Humphreys, P. & Berkeley, D. (1985). Handling uncertainty- levels of analysis of decision problems. In G. Wright (Ed.), *Behavioral decision making* (pgs 257-282). Plenum Press; New York.

Janis, I. L. (1989). Crucial decisions: Leadership in policymaking and crisis management. New York: The Free Press.

Kanki, B. G., (1996). Stress and aircrew performance: A team level perspective. In G. Driskell and E. Salas (Eds.), *Stress and human performance* (pp. 127- 162). Mahwah, NJ: Lawrence Erlbaum Associates.

Klein, G. A. (1983). Utility and decision strategies: A second look at the rational decision maker. In *Organizational Behavior and Human Performance*, 31, 1-25.

Klein, G. A., Calderwood, R., & Clinton Cirocco, A. (1986). Rapid decision making on the fireground. In *Proceedings of the Human Factors Society 30th Annual Meeting* (pp. 576-580). Santa Monica, CA: Human Factors and Ergonomics Society.

Klein, G. A. (1989). Recognition-primed decisions. In W. B. Rose (Ed.), *Advances in man-machine systems research* (Vol. 5, pp.47-92). Greenwich, CT: JAI.

Klein, G. A. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 138-147). Norwood, NJ: Ablex.

Klein, G. A. (1997a). The status of the naturalistic decision making framework. In R. Flin, E. Salas, M. Strub, & L. Martin (Eds.), *Decision making under stress* (pp. 11-28). Brookfield, VT: Ashgate.

Klein, G. A. (1997b). An overview of naturalistic decision making applications. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic decision making* (pp. 49-59). Mahwah, NJ: Lawrence Erlbaum Associates.

Klein, G. A. (1998). Sources of power: How people make decisions. Cambridge, MA: MIT Press.

Klein, G.A. (2000). How can we train pilots to make better decisions? In H. F. O'Neil, & D. H. Andrews (Eds.), *Aircrew training and assessment*, (pp.165-195). Mahwah, NJ: LEA.

Kleinman, D. L., Luh, P. B., Pattipati, K. R., & Serfaty, D. (1992). Mathematical models of team performance: A distributed decision-making approach. In W. S. Swezey & E. Salas, (Eds.), *Teams: their training and performance* (pp. 177-217). Norwood, NJ: Ablex Publishing.

Kuperman, G. G. (1998). Cognitive systems engineering for battlespace dominance. Paper presented at the Fourth Conference on Naturalistic Decision Making, Warrenton, VI.

Larkin, J. H. (1983). The role of problem representation in physics. In D. Gentner & A. L. Stevens (Eds.) *Mental models*. Hillsdale, NJ: Erlbaum.

Leedom, D., Adelman, L., & Murphy, J. (1998). *Critical Indicators in Naturalistic Decision Making*. Paper presented at the Fourth Conference on Naturalistic Decision Making, Warrenton, VI.

Lipshitz, R. (1995). Converging themes in the study of decision making in realistic settings. In G Klein, J. Orasanu, R. Calderwood, & C. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 103-137). Norwood, NJ: Ablex.

Maule, A. J. (1985). Cognitive approaches to decision making. In G. Wright (Ed.) *Behavioral decision making* (pp.61-84). New York-Plenum Press.

Metcalfe, J., & Shimamura, A.P. (Eds.). (1996). *Metacognition: Knowing about knowing*. Cambridge, MA: MIT Press.

Means, B, Crandall, B., Salas, E., & Jacobs, T. O. (1993). Training decision makers for the real world. In G. Klein, J. Orasanu, R. Calderwood, and C. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 306-326). Norwood, NJ: Ablex.

Naval Safety Center. (1997, October). *Human Factors Accident Analysis Classification System* (OPNAV 3750.6R, Appendix O). Norfolk, VI.

Nullmeyer, R. A., & Spiker, V. A. (1999). The importance of crew resource management in MC-130P mission performance: Implications for training effectiveness evaluation. *Proceedings of the 21st Interservice/Industry Training Systems and Education Conference.* Orlando, FL.

Orasanu, J. M. (1993). How do we recognize a good decision in the cockpit? In E. Weiner, B. Kanki, & R.L. Helmreich (Eds.), *Cockpit resource management* (pp. 150-153). San Diego, CA: Academic Press.

Orasanu, J. M. (1997). Stress and naturalistic decision making: Strengthening the weak links. In Flin, R., Salas E., Strub, M., & Martin, L. (Eds.), *Decision making under stress: emerging themes and applications* (pp. 43-66). Brookfield, VT: Ashgate.

Orasanu, J., & Connolly, T. (1993). The reinvention of decision making. In G. A. Klein, J. Orasanu, R. Calderwood. & C. E. Zsambock (Eds.), *Decision making in action: Models and methods* (pp. 3-20). Norwood, NJ: Ablex.

Orasanu, J., Dismukes, K., & Fischer, U. (1993). Decision errors in the cockpit. In *Proceedings of the human Factors and Ergonomics Society 37th Annual Meeting*, 1, 363-367. Santa Monica, CA: Human Factors and Ergonomics Society

Orasanu, J., & Fischer, U. (1992). Team cognition in the cockpit: Linguistic control of shared problem solving. In *Proceedings of the 14th Annual Conference of the Cognitive Sciences Society*, (pp 189-194). Hillsdale, NJ: LEA.

Orasanu, J., & Fischer U. (1997). Finding decisions in natural environments: The view from the cockpit. In C. Zsambok & G. Klein (Eds.), *Naturalistic decision making.* (pp.343-357). Hillsdale, NJ: LEA.

Orasanu, K. J. (1999). Shared problem models and flight crew performance. In *Aviation Psychology and Practice* (pp. 225-285). Brookfield, VT: Ashgate.

Patel, U. L., & Groen, G. J. (1991). The general and specific nature of medical expertise: A critical look. In K. A. Ericson, and J. Smith (Eds.), *Towards a general theory of expertise: Prospects and limits*, (pp. 93-125), New York: Cambridge University Press.

Pascual, R. & Henderson, S. (1997). Evidence of naturalistic decision making in military command and control. In C.E. Zsambok & G. Klein (Eds.), *Naturalistic decision making* (pp. 217-226). Mahwah, NJ: Erlbaum.

Payne, J. W. (1982). Contingent decision behavior. *Psychological Bulletin*, 92, 382-402.

Payne, J. W. (1985). Psychology of risky decisions. In G. Wright (Ed.), *Behavioral decision making* (pp. 3-23). New York-Plenum Press.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology Learning, Memory and Cognition,* 14, 534-552.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). The adaptive decision maker. New York: Cambridge University Press.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1990). The adaptive decision maker: Effort and accuracy in choice. In R. M Hogarth (Ed.), *Insights in decision making: a tribute to Hillel J. Einhorn* (pp. 129-153). Chicago: University of Chicago Press.

Pounds, J., & Fallesen, J. (1995, November). *Familiarity Effects on Strategy Use in Tactical Problem Solving.* Poster session presented at the Conference of the Society for Judgment and Decision Making, Los Angeles, CA.

Prince. C., Hartel, C., & Salas, E. (1992). Aeronautical decision making and consistency of crew behaviors: Implications for training. *Proceedings of the Seventh International Symposium of Aviation Psychology,* (pp.248-251). Columbus, OH.

Reason, J. R. (1990). *Human error.* Cambridge, UK: Cambridge University Press.

Reinartz, S. J. (1993). An empirical study of team behaviour in a complex and dynamic problem-solving context: A discussion of methodological and analytical aspects. Gruppe Arbeitsschultz und Ergonomie. Cologne: TUV Rheinland.

Rentsch, J. R., Heffner, T. S., & Duffy, L.T. (1994). What you know is what you get from experience: Team experience related to teamwork schemas. *Group & Organizational Management,* 19, 450-474.

Rouse, W. B., & Valusek, J. (1993). Evolutionary design of systems to support decision making. In C. Zsambok & G. A. Klein (Eds.), *Decision making in action: Models and methods* (pp. 270-286). Mahwah, NJ: LEA.

Sarter, N. B., & Woods, D. D. (1997). Teamplay with a powerful and independent agent: A corpus of operational experiences and automation surprises on the airbus A-320. *Human Factors*, 39 (4) 553-569.

Seamster, T. L., Redding, R.E., & Kaempf, G. L. (1997). *Applied Cognitive Analysis*. Brookfield, VT: Ashgate.

Shanteau, J. (1986). Psychological characteristics of expert decision makers. Paper presented at the University of Southern California Symposium on Audit Judgment and Expert Systems, Los Angeles, February.

Shanteau, J. (1988). Psychological characteristics and strategies of expert decision makers. In *Advances in decision research*. B. Rohrmann, L. R. Beach, C. Vlek, & S. R. Watson (Eds.), Amsterdam: North Holland.

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, 69, 99-118.

Simon, H. A. (1979). Information processing models of cognition. *Annual Review of Psychology*, 30, 363-396.

Slovic, P, Fischoff, B., & Lichtenstein, S. (1982). Facts versus fears: Understanding perceived risk. In D. Kahneman, P. Slovis, & A, Tversky (Eds.) *Judgement under uncertainty: Heuristics and biases* (pp. 463-489). Cambridge, UK: Cambridge University Press.

Smith-Jentsch, K. A., Johnston, J. H., & Payne, S. C. (1998). Measuring team-related expertise in complex environments. In J. A. Cannon-Bowers & E. Salas (Eds.), Making decisions under stress: Implications for individual and team training (pp.61-87). Washington, DC: APA Press.

Stokes, A. F., Kemper, K., & Kite, K. (1997). Aeronautical decision making, cue recognition, and expertise under time pressure. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic decision making* (pp. 183-196). Mahwah, NJ: Erlbaum.

Spiker, V.A., Tourville, S. J., Silverman, D.R., & Nullmeyer, R. T. (1996, November). *Team performance during combat mission training: A conceptual model and measurement framework* (AL/HR-TR-1996-0092). Mesa, AZ: Aircrew Training Research Division.

Spiker, V. A., Silverman, D. R., Tourville, S. J., & Nullmeyer, R. T. (1998). Tactical team resource management effects on combat mission training performance. (AL/HR-TR-1997-0137). Mesa, AZ: Air Force Research Laboratory Warfighter Training Division.

Spiker, V. A., Tourville, S. J., Bragger, J. B., Dowdy, D., & Nullmeyer, R. L. (1999). Measuring C-5 crew coordination proficiency in an operational wing. In *Proceedings of the 21st Interservice/Industry Training Systems and Education Conference*. Orlando, FL.

Tashakkori, A., & Teddlie, C. (1998). Mixed methodology- combining qualitative and quantitative approaches. (Applied social research methods; v. 46). Thousand Oaks, CA: Sage Publications.

Tetlock, P. E. (1985). Accountability: The neglected social context of judgment and choice. *Research in Organizational Behavior*, 7, 297-332.

Thompson, J. S., Tourville, S. J., Spiker, V.A., & Nullmeyer, R. T. (1999). Crew resource management and mission performance during MH-53J combat mission training. In *Proceedings of the 20th Interservice/Industry Training Systems and Education Conference*. Orlando, FL.

Thorton, R. C., Kaempf, G. L., Zeller, J. L., & McAnuity, D. M. (1992), July). *An evaluation of crew coordination and performance during simulated UH-60 helicopter mission* (ARI Research Note 92-63). Ft Rucker, AL: United States Army Research Institute for the behavioral and Social Sciences.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-232.

Zsambok, C. E. (1993). Advanced team decision making in C^2 settings. *Proceedings of the 1993 Symposium on Command and Control Research*, 45-52. McLean, VA:SAIC.

Zsambok, C. E. (1997). Naturalistic decision making: Where are we now? In C. E. Zsambok & G. E. Klein (Eds.), *Naturalistic decision making_* (pp.3-16). Mahwah, NJ: LEA.

APPENDIX A

PARTICIPANT DEMOGRAPHICS

PROCESS SCORE RATER DEMOGRAPHICS

Table A1
Participant Demographics

Crew 1	**PILOT**	**COTAC**	**Crew Total**	**Crew Mean**
Age	25	25		25
Gender	M	M	2M	
Crewed Events Together			4	
Days Since Last Flight	1	12		
Days Since Last Sim	3	1		
S-3B Flight Hours	45	45	**90**	**45**
Other Flight Hours	370	145	515	257.5
Combined Flight Hours	415	190	605	302.5
S-3B Sim Hours	35	35	70	35
Combined S-3 Sim and Flight Hrs	80	80	160	80

Crew 2	**Pilot**	**COTAC**	**Crew Total**	**Crew Mean**
Age	24	25		24.5
Gender	M	M	2M	
Crewed Events Together			8	
Days Since Last Flight	21	2		
Days Since Last Sim	3	1		
S-3B Flight Hours	38	50	**88**	**44**
Other Flight Hours	532	180	712	356
Combined Flight Hours	570	230	800	400
S-3B Sim Hours	62	80	142	71
Combined S-3 Sim And Flight Hrs	100	130	230	115

Crew 3	**Pilot**	**COTAC**	**Crew Total**	**Crew Mean**
Age	29	28		28.5
Gender	M	M	2M	
Crewed Events Together			1	
Days Since Last Flight	8	2		
Days Since Last Sim	11	7		
S-3B Flight Hours	12	40	**52**	**26**
Other Flight Hours	350	146	496	248
Combined Flight Hours	362	186	548	274
S-3B Sim Hours	45	80	125	62.5
Combined S-3 Sim And Flight Hrs	54	120	174	87

Crew 4	**Pilot**	**COTAC**	**Crew Total**	**Crew Mean**
Age	27	24		25.5
Gender	M	M	2M	
Crewed Events Together			1	
Days Since Last Flight	7	16		
Days Since Last Sim	1	13		
S-3B Flight Hours	40	30	**70**	**35**
Other Flight Hours	260	130	390	195
Combined Flight Hours	300	160	460	230
S-3B Sim Hours	30	30	60	30
Combined S-3 Sim And Flight Hrs	70	60	130	65

(table continues)

Crew 5	**Pilot**	**COTAC**	**Crew Total**	**Crew Mean**
Age	26	26		26
Gender	M	M	2M	
Crewed Events Together			3	
Days Since Last Flight	26	24		
Days Since Last Sim	9	16		
S-3B Flight Hours	105	78	**183**	**91.5**
Other Flight Hours	245	100	345	172.5
Combined Flight Hours	350	178	528	264
S-3B Sim Hours	75	73	148	74
Combined S-3 Sim And Flight Hrs	180	151	331	165.5

Crew 6	**Pilot**	**COTAC**	**Crew Total**	**Crew Mean**
				26
Age	25	27		
Gender	M	M	2M	
Crewed Events Together	0	0	0	
Days Since Last Flight	4	23		
Days Since Last Sim	3	9		
S-3B Flight Hours	80	100	**180**	**90**
Other Flight Hours	275	100	375	187.5
Combined Flight Hours	355	200	555	277.5
S-3B Sim Hours	100	100	200	100
Combined S-3 Sim And Flight Hrs	180	200	380	190

Crew 7	**Pilot**	**COTAC**	**Crew Total**	**Crew Mean**
Age	27	24		25.5
Gender	M	M	2M	
Crewed Events Together			0	
Days Since Last Flight	26	4		
Days Since Last Sim	6	0		
S-3B Flight Hours	80	50	**130**	**65**
Other Flight Hours	378	100	478	239
Combined Flight Hours	458	150	608	304
S-3B Sim Hours	100	100	200	100
Combined S-3B Sim And Flight Hrs	180	150	330	165

Crew 8	**Pilot**	**COTAC**	**Crew Total**	**Crew Mean**
Age	26	24		25
Gender	M	M	2M	
Crewed Events Together			4	
Days Since Last Flight	10	3	13	
Days Since Last Sim	14	17	31	
S-3B Flight Hours	30	24	**54**	**27**
Other Flight Hours	340	110	450	225
Combined Flight Hours	370	134	504	252
S-3B Sim Hours	40	40	80	40
Combined S-3B Sim And Flight Hrs	70	64	134	67

(table continues)

Crew 9	**Pilot**	**COTAC**	**Crew Total**	**Crew Mean**
Age	27	34		30.5
Gender	M	M	2M	
Crewed Events Together			20	
Days Since Last Flight	4	13		
Days Since Last Sim	3	5		
S-3B Flight Hours	950	2100	**3050**	**1525**
Other Flight Hours	250	400	650	325
Combined Flight Hours	1200	2500	3700	1850
S-3B Sim Hours	100*	120*	220	110
Combined S-3 Sim And Flight Hrs	1050	2220	3270	1635

Crew 10	**Pilot**	**COTAC**	**Crew Total**	**Crew Mean**
Age	36	40		38
Gender	M	M	2M	
Crewed Events Together			2	
Days Since Last Flight	0	337		
Days Since Last Sim	127	338		232.5
S-3A/B Flight Hours	2200	2000	**4200**	**2100**
Other Flight Hours	1600	100	1700	850
Combined Flight Hours	3800	2100	5900	2950
S-3A/B Sim Hours	250*	120*	370	185
Combined S-3 Sim And Flight Hrs	2450	2120	4570	2285

Crew 11	**Pilot 11**	**COTAC 11**	**Crew 11 Total**	**Crew 11 Mean**
Age	30	28		29
Gender	M	M	2M	
Crewed Events Together			8	
Days Since Last Flight	5	11		
Days Since Last Sim	15	36		
S-3B Flight Hours	1000	570	**1570**	**785**
Other Flight Hours	240	152	392	196
Combined Flight Hours	1240	722	1962	981
S-3B Sim Hours	100	100	200	100
Combined S-3 Sim And Flight Hrs	1100	670	1770	885

Crew 12	**Pilot 12**	**COTAC 12**	**Crew 12 Total**	**Crew 12 Mean**
Age	28	27		27.5
Gender	M	M	2M	
Crewed Events Together			5	
Days Since Last Flight	4	4		
Days Since Last Sim	6	26		
S-3B Flight Hours	900	730	1630	**815**
Other Flight Hours	265	160	425	212.5
Combined Flight Hours	1165	890	2055	1027.5
S-3B Sim Hours	85	150	235	117.5
Combined S-3B Sim And Flight Hrs	985	880	1865	932.5

(table continues)

Crew 13	PILOT 13	COTAC 13	CREW 13 TOTAL	CREW 13 MEAN
Age	41	34		37.5
Gender	M	M	2M	
Crewed Events Together			5	
Days Since Last Flight	2	15		
Days Since Last Sim	210	19		
S-3A/B Flight Hours	2500	2100	**4600**	**2300**
Other Flight Hours	1340	280	1620	810
Combined Flight Hours	3840	2380	6220	3110
S-3A/B Sim Hours	300	400	700	350
Combined S-3 Sim And Flight Hrs	2800	2500	5300	2650

Crew 14	Pilot 14	COTAC 14	Crew Total	Crew Mean
Age	32	27		29.5
Gender	M	M	2M	
Crewed Events Together			15	
Days Since Last Flight	2	1		
Days Since Last Sim	7	7		7
S-3B Flight Hours	650	950	**1600**	**800**
Other Flight Hours	2150	90	2240	1120
Combined Flight Hours	2800	1040	3840	1920
S-3B Sim Hours	150	120	270	135
Combined S-3 Sim And Flight Hrs	800	125	925	462.5

Crew 15	Pilot	COTAC	Crew Total	Crew Mean
Age	36	33		34.5
Gender	M	M	2M	
Crewed Events Together			25	
Days Since Last Flight	0	55		27.5
Days Since Last Sim	14	76		45
S-3B Flight Hours	650	1000	**1650**	**825**
Other Flight Hours	290	150	440	220
Combined Flight Hours	940	1150	2090	1045
S-3B Sim Hours	150	125	275	137.5
Combined S-3 Sim And Flight Hrs	800	1125	1925	962.5

Crew 16	Pilot	COTAC	Crew Total	Crew Mean
Age	41	38		39.5
Gender	M	M	2M	
Crewed Events Together			0	
Days Since Last Flight	385	4		194.5
Days Since Last Sim	1342	125		733
S-3B Flight Hours	3700	1950	**5650**	**2825**
Other Flight Hours	300	120	420	210
Combined Flight Hours	4000	2070	6070	3035
S-3B Sim Hours	200	200	400	200
Combined S-3 Sim And Flight Hrs	3900	2150	6050	3025

Table A2
Process Score Rater Demographics

Rater	Designator// Affiliation	Age	Years as S-3 Instructor	S-3 Flight Hours	Other Flight Hours
1	Pilot/Civ	59	7.5	1900	2100
2	Pilot/Civ	54	4	1100	4000
3	Pilot/Civ	54	5	1100	17,000
4	NFO/Civ	36	3	1200	750
5	NFO/Civ	36	4	2350	250
6	NFO/Civ	40	11	2100	7000
7	NFO/Civ	39	9	1005	1400
8	NFO/Mil	38	7	2760	500

APPENDIX B

APPROVAL TO CONDUCT RESEARCH WITH NAVAL AVIATORS

APPROVAL TO CONDUCT RESEARCH WITH NAVAL AVIATORS ASSIGNED TO VS-41 AND SEA CONTROL WING, U.S. PACIFIC FLEET SQUADRONS

DEPARTMENT OF THE NAVY
COMMANDER NAVAL AIR FORCE
UNITED STATES PACIFIC FLEET
P.O. BOX 357051
SAN DIEGO, CALIFORNIA 92135-7051

1542
Ser N45/ 683
MAY 22 2000

University of San Diego
Office of the Provost
5998 Alcala Park
San Diego, CA 92110

Attention: Human Subjects Committee

Ladies and Gentlemen:

Ms. Constance Gillan has discussed with me her proposal to conduct flight-training research involving North Island based personnel under my command's cognizance. Her proposed comparative study of novice versus expert decision-making and situational awareness will require full access to S-3 aircraft flight simulator data as well as personal interviews and observation of involved subjects. I fully support this endeavor.

Research of this precise nature into cognitive aircrew processes in the time-critical cockpit environment is sorely needed. As a member of the Naval Aviation Human Factors Quality Management Board since 1997, I have closely monitored Department of Defense, academic and commercial efforts in this area as part of our charter to significantly reduce human-error aviation mishaps. Ms. Gillan's proposed research could provide data and conclusions directly applicable to that effort.

I stand by to enable and assist Ms. Gillan in her research in every way possible. For questions, I may be contacted at (619) 545-2788, or e-mail to keepper.robert.h@cnap.navy.mil.

Sincerely,

R. H. KEEPPER
Captain, U.S. Navy
Force Safety Officer
By direction of the Commander

Copy to:
COMSEACONWINGPAC (N013)

DEPARTMENT OF THE NAVY
COMMANDER, SEA CONTROL WING, U.S. PACIFIC FLEET (92135-7131)
COMMANDING OFFICER, SEA CONTROL SQUADRON FOUR ONE (92135-7098)
NAVAL AIR STATION, NORTH ISLAND, CA

1542	1542
Ser N00/ 0444	Ser N00/ 108
21 JUNE 2000	21 JUNE 2000

University of San Diego
Office of the Provost
Attn: Human Subjects Committee
5998 Alcala Park
San Diego, CA 92110

Gentlemen,

Ms. Constance Gillian has approached the Commanding Officer of Sea Control Squadron FOUR ONE and me about conducting flight-training research using students at the Sea Control Squadron FOUR ONE training command and fleet aviators assigned to Sea Control operational fleet squadrons located at Naval Air Station North Island. She has proposed to collect observational data related to situational awareness and decision making from novice and more experienced flight crew performance in a full flight simulator scenario followed by debrief interviews with study participants

Both the Commanding Officer of Sea Control Squadron FOUR ONE and I have reviewed Ms. Gillan's doctoral dissertation study proposal. We believe that her research will provide significant benefit to the advancement of human factors initiatives and conduct of training both at Sea Control Squadron FOUR ONE and throughout naval aviation training. We fully support her research effort and will stand by to assist her in any way we can. If you have any questions, please don't hesitate to call me at (619) 545-5259

Sincerely,

G. PALABUDA
Commander, U. S. Navy
Commanding Officer

Sincerely,

D. E. HEPFER II
Captain, U. S. Navy
Commander

APPENDIX C

PARTICIPANT CONSENT FORM

Sample Participant Consent Form

Constance (Connie) Gillan is conducting research toward a doctorate in Leadership Studies at the University of San Diego under the direction of Dr. Mary Woods Scherr. You have been asked to participate in this study because you are in a full-time training status or a fleet aviator. Connie is conducting an investigation into how aircrews with varying levels of experience in the same aircraft model may differ in their use of decision-making strategies and how they process decisions.

Your participation will involve permitting Connie and two instructors to remain in the event brief and debrief as well as observe and evaluate your actions in a 20-minute simulator event from the instructor console. The researcher and instructors will make video and audio recordings using the Computer Assisted Debriefing System (CADS). The researcher and instructors may also make written notes during the simulator event and associated briefs, debriefs, and interviews with your crew as necessary and may photocopy the evaluation sheet for the observed event. To ensure you remain anonymous no names will be used on the observation/evaluation sheet or anywhere in the data collection, analysis or final research paper. Your crew will have a code associated with it for all purposes of the study data collection, analysis, and reporting/publishing of the study. You may view all notes and evaluations associated with your study event as well as the transcripts of the debrief/interview. No one other than the researcher and observer/evaluators will have access to the raw data. There may be persons outside the command that will assist in transcribing the interviews and compiling the raw scores for analysis.

Participation in this study is voluntary and data collected will be used for training purposes only. You may refuse to participate or withdraw at any time without penalty. You may request that video and digital recording be stopped at any time and files/tapes erased. You may refuse to have files/tapes viewed by other persons other than the researcher and instructor(s) for your event. Approximately one hour of additional time beyond the scheduled event time will be required to respond to a series of open-ended questions about your situational assessment and decision making processes in the simulator event

There may be no direct benefit to you from these procedures although you may gain experience and feedback in a scenario that you may or may not have been exposed to previously. The results of this study may help in the advancement of decision making research, design of aircrew training and operational procedures in aviation.

If you choose to participate in the scenario you must pledge not to discuss it with other crews until informed that the data collection and analysis phases of this study are complete so the integrity of the study is not compromised.

You may call the University of San Diego Human Subjects Committee Office at (619) 260-6889 to inquire about your rights as a research subject and/or report research related problems to your Commanding Officer.

Connie has explained this study to you and answered your questions. If you have other questions or research related problems, you may reach Connie at 545-1823 or send her e-mail at cgillan@adnc.com. Research records will be kept anonymous and confidential and will be destroyed after three years (requirement for research purposes).

You have received a copy of this consent document to keep.

I, the undersigned, understand the above explanations and on that basis, I give consent to my voluntary participation in this research.

Signature of Subject Date
Naval Air Station North Island, San Diego, CA

Signature of Principal Researcher Date

Signature of Witness Date

APPENDIX D

SCENARIO SYNOPSIS WITH NATURALISTIC DECISION MAKING ELEMENTS

Scenario Synopsis and Naturalistic Decision Making Elements for each Event

Event 1: Takeoff and Departure

Synopsis: No abnormal indications were presented during this segment of the scenario.

Naturalistic Decision Making Elements:

Ill-structured problems	No
Uncertain dynamic environments	No
Shifting, ill-defined, or competing goals	No
Action / feedback loops	Yes
Time stress	Low
High stakes	Low
Multiple players	Yes (internal)
Organizational goals and norms	Yes

Difficulty Level: Low

Cues: Normal Take-off and Departure

Appropriate Response: Perform in accordance with normal procedures and operations.

Event 2: Starter Light No.1 Engine

Synopsis: Significant safety of flight related event – flight manual procedures require the engine to be secured and for the crew to return for a landing while exercising single engine approach and recovery procedures. Depending upon the circumstances, it would not be unusual for a crew to declare an in-flight emergency if the situation deteriorated or became more complicated than a simple single engine approach.

Naturalistic Decision Making Elements:

Ill-structured problems	No – well defined procedures
Uncertain dynamic environments	Middling
Shifting, ill-defined, or competing goals	No – checklists and recovery
Action / feedback loops	Yes
Time stress	High
High stakes	Middling
Multiple players	Yes (internal)
Organizational goals and norms	Yes

Difficulty Level: Low

Cues: The flashing red Master Warning Light which is located in an exceedingly prominent spot on the instrument glare shield (directly in front of each crewmember's seat and dead center in their forward line of view) will illuminate as will the slightly less commanding Starter Caution Light on the systems annunciator panel.

Appropriate Response:

Procedural:	Immediately execute the "boldface" steps (those procedures which have been committed to memory) associated with the Starter Caution Light In-Flight procedures. Then complete the remaining steps of the Starter Caution Light In-Flight emergency procedure using the Pocket Checklist (PCL) as a reference.
Tactical:	Initiate a turn towards a suitable airport
Strategic:	Start planning for a single engine recovery

Event 3: Checklist Interruption (Other aircraft inbound)

Synopsis: While the crew was engaged in completing a checklist associated with the Starter Caution Light procedure they were informed of another aircraft proceeding to the same airfield with its own emergency.

Naturalistic DM Elements:

Ill-structured problems	No
Uncertain dynamic environments	Yes
Shifting, ill-defined, or competing goals	Yes
Action / feedback loops	Yes
Time stress	Moderate
High stakes	Yes
Multiple players	Yes (internal and external)
Organizational goals and norms	Yes

Difficulty Level: Relatively straightforward and simple

Cues: The information associated with another aircraft inbound to the field with the potential to cause a delayed recovery for the scenario crew was very clearly provided by Air Traffic Control (ATC). The cues associated with particulars of the emergency declared by the crew of the other aircraft in close vicinity were curtly communicated by ATC.

Appropriate Response:

Procedural: Acknowledge the information related by ATC that references the other aircraft with an emergency inbound.

Tactical: Crews should announce their intentions to continue inbound for an arrested landing in advance of the other emergency aircraft.

Strategic: Crews should-emphasize their desire and intentions to get on the deck as soon as possible without becoming distracted by the other emergency aircraft.

Event 4: Low Oil No. 2 Engine

Synopsis: significant safety of flight related event on par with a Starter Caution Light in-flight. The biggest difference between the two Emergency Procedures was the issue of immediacy in terms of how quickly the offending engine would need to be secured. In the case of the Starter Caution Light in flight, continued operation of the engine with the light illuminated could result in a catastrophic and potentially explosive engine failure. In the case of the Low Oil pressure indication the Emergency Procedures require the crew to retard the engine's throttle to "IDLE" and then wait to see how the engine responds. If the Low Oil indication persists, the procedure requires the engine to be secured (shut down) to preclude the engine from seizing which would most likely cause an engine fire and quite possibly a catastrophic failure of associated systems.

In this scenario, the crew is presented with a very real dilemma. With one engine secured, the remaining engine presents a condition that would normally require the only remaining engine to be secured as well. (Note the S-3B is incapable of gliding without engine thrust. Furthermore, NATOPS specifically prohibits S-3 crews from attempting un-powered landings following a dual engine failure/flameout.)

Naturalistic Decision Making Elements:

Ill-structured problems	Yes
Uncertain dynamic environments	Yes
Shifting, ill-defined, or competing goals	Yes
Action / feedback loops	Multiple and conflicting
Time stress	Yes, significant
High stakes	Yes
Multiple players	Yes (internal and external)
Organizational goals and norms	Yes

Difficulty level: Simple; only two options (either secure the failing engine or not), but a greatly complicated decision making process due to the unprecedented nature of the compounded emergencies and the lack of documented decision making guidance associated with this type of challenge.

Cues: Under normal circumstances, the Master Caution Light would illuminate and then flash as the engine oil pressure decreased below a predetermined level. However, in this particular scenario there would have been no apparent "attention getting" warnings (no flashing lights, no aural warnings, etc.) associated with the decreasing oil pressure. Instead, the only cues available to the crew would have been an obscure indication of decreasing engine oil pressure on a small, less than prominent, quarter-size, analog style engine oil pressure gage.

Appropriate Response:

Procedural: In this scenario, the crew would be force to apply a line of rationale in apparent conflict with the "standard" emergency procedures (i.e., even though the procedures for an isolated Low Oil Pressure indication require the engine to be secured; in this instance, a deviation from the procedure was necessary to keep at least one operating engine on line.)

Tactical: Expedite the recovery process by flying in the most direct manner to the nearest airfield capable of recovering an S-3. The significant caveat being that consideration had to be given to the intended flight path and the risks associated with the pending loss of the only remaining engine (i.e., if the engine quit prior to reaching the field would the aircraft be in a position to inflict the least amount of collateral damage to personnel and property on the ground?)

Strategic: Reprioritize any game plans to address the increasingly likely failure of the only remaining engine. If the crew had not previously declared an In-Flight Emergency, it would have been appropriate to do so immediately following the crew's comprehension of their Low Oil Pressure condition.

Event 5: Approach Priorities

Synopsis: In anticipation of both a single engine approach and its inherent potential for a single engine wave-off crews would normally dump fuel in order to both reduce their gross weight (so as to not exceed the arresting gear limitations) and to improve their single engine climb capabilities. Less weight implies more excess thrust available to improve the climb gradient. Crews should have given consideration to contingencies associated with single engine approach and go-around profiles.

Naturalistic Decision Making Elements:

Ill-structured problems	Yes
Uncertain dynamic environments	Yes
Shifting, ill-defined, or competing goals	No
Action / feedback loops	Yes but poorly defined
Time stress	High
High stakes	Yes and increasing
Multiple players	Yes (internal and external)
Organizational goals and norms	Yes

Difficulty Level: Moderately difficult and increasingly difficult as the scenario progressed

Cues: In its capacity as a decision making guide the NATOPS Pocket Checklist (PCL) specifically states that fuel should be dumped "as required." When confronted with the possibility of a single engine approach and wave-off with only one marginally operative engine crews should have recognized the need to reduce their gross weight by dumping fuel.

Other cues such as the deteriorating condition of the No. 2 engine, the likelihood that the aircraft would be unable to safely clear Point Loma in the event of a wave-off and the need to determine the resultant direction of turnout following the wave-off should have all contributed the crew conducting contingency planning.

Appropriate Response:

Procedural: Crews should dump fuel.

Tactical: Crews should request an approach type appropriate to their level of extremes. A visual straight-in approach or vectors to the initial approach point would most likely result in the most effective aircraft positioning.

Crews should brief hook skip contingencies and wave-off techniques. Specifically, rudder application, rate of throttle movement, and direction of turn following wave-off initiation.

Strategic: Crews should declare an emergency and requested priority handling from Air Traffic Control (ATC).

Event 6: Game Plan for Single Engine Recovery without Field Arresting Gear

Synopsis: Due to the overall deterioration in the aircraft's mechanical condition and the crew's selection of a landing field located in close proximity to a major metropolitan area contingency planning should have been discussed between the pilot and COTAC.

Naturalistic Decision Making Elements:

Ill-structured problems	Yes
Uncertain dynamic environments	Yes
Shifting, ill-defined, or competing goals	Yes
Action / feedback loops	Yes
Time stress	Higher
High stakes	Yes
Multiple players	Yes (internal and external)
Organizational goals and norms	Conflicted

Difficulty Level: Difficult, given the number of items either occupying or outright demanding the crew's attention finding time for contingency planning would have become exponentially more difficult as the scenario progressed.

Cues: The overall extremis of the situation should have provided the most unobservant of crews with a general idea that they should be developing a "Plan B".

Tower informed crew that arresting gear was unavailable; crews should have recognized this as a departure from the normal emergency profile.

Tower informed crew that the landing environment was fouled; this should have triggered an immediate discussion as to where they intended to position the jet in preparation for an imminent initiation of the ejection sequence.

Appropriate Response:

Procedural: Continue to adjust the approach profile in order to accommodate the ever-shrinking runway availability.

Crews should discuss their emergency egress criteria once it became apparent that they would be landing in a configuration with a large potential for a runway departure (or excursion) during the landing roll out.

Tactical: Continue to exercise all available options and discuss unacceptable safety of flight excursions that would necessitate an ejection.

Strategic: Maintain awareness of the "big picture" by communicating intentions both internally to each other and externally to Air Traffic Control (ATC).

Event 7: Foul Deck Final Decision

Synopsis: Under normal conditions, any sort of fouled deck condition would necessitate a wave-off even under single engine conditions. In this scenario, the opportunity to execute a single engine wave-off was negated due to the progressively worsening condition of the only operative motor. In other words, given the engine's low oil condition the probability of inducing an engine failure due to the low oil condition by advancing the throttle, as was required by executing a single-engine go-around, was a very real likelihood.

Naturalistic Decision Making Elements:

Ill-structured problems	Yes
Uncertain dynamic environments	Yes
Shifting, ill-defined, or competing goals	Yes
Action / feedback loops	Yes
Time stress	Extreme
High stakes	Yes
Multiple players	Yes (internal and external)
Organizational goals and norms	Unclear

Difficulty Level: Extremely difficult

Cues: Verbal communication from Tower. Conceivably there was the possibility of this being a somewhat ambiguous cue due to a very terse report from the Tower during a period of high workload for the crew (i.e., presence of other factors competing for the crews' attention such as aircraft control, awareness or monitoring of the dying engine, amount of mental faculties devoted to contingency planning, etc.)

Appropriate Response:

Procedural: Either continue with the landing or initiate a wave-off.

Tactical: The generally accepted "approved response" for this scenario was for the crew to either: (a) disregard the normal implications associated with a fouled deck (i.e., wave-off) and modify their profile and land anyway while attempting to stop well short of the intersection or (b) wave-off and purposefully position the aircraft over the runway in a safe ejection envelope.

Strategic: Scenario end game.

APPENDIX E

SIMULATOR DATA SETS USED FOR OUTCOME RANKING

Outcome Data

FINAL SNAPSHOT	CREW 1 EJECT	CREW 2 EJECT	CREW 3 EJECT	CREW 4 EJECT	CREW 5 LAND	CREW 6 EJECT	CREW 7 EJECT	CREW 8 EJECT	CREW 9 LAND	CREW10 LAND	CREW11 LAND	CREW 12 EJECT	CREW13 LAND	CREW14 LAND	CREW15 LAND	CREW16 LAND
Airspeed	110	120	115	115	#	90	115	130	#	#	#	115	#	#	#	#
Altitude	100	50	90	200	#	0	500	350	#	#	#	500	#	#	#	#
Angle of Attack	17	14	16	15	#	10	15	^	#	#	#	15	#	#	#	#
Angle of bank	0	0	5	0	#	0	0	5	#	#	#	0	#	#	#	#
Flap Position	Takeoff	Takeoff	Takeoff	Takeoff	Takeoff	Takeoff	Takeoff	Takeoff	Takeoff	Takeoff	Takeoff	Takeoff	Takeoff	Takeoff	Takeoff	Takeoff
Gear Position	Down	Down	Down	Down	Down	Down	Down	Up	Down	Down	Down	Up	Down	Down	Down	Down
Tailhook Position	Down	Down	Down	Down	Down	Down	Down	Down	Down	Down	Down	Up	Down	Down	Down	Down
Fuel Quantity (lbs)	10.3	10	10.3	10	10.3	5	8	3	9.7	6.6	7.7	8.7	6.9	8	8.1	7.8
Speedbrake Position	IN	IN	IN	IN	OUT	IN	IN	IN	OUT	IN	IN	IN	•	OUT	OUT	OUT
Emergency Hyd Pump	ON	ON	ON	ON	ON	ON	ON	ON	ON	ON	ON	ON	ON	ON	ON	ON
Vertical Speed Indicator	100 Down	0	800 Down	1000 Down	#	0	1000 Down	^	#	#	#	1500 Down	#	#	#	#
Compass Heading	290	290	180	290	290	290	290	170	#	#	#	360	#	#	#	#
Geographical Position	FIELD	FIELD	Over water south	FIELD	FIELD	Runway 29	FIELD	Field/head to water	#	#	#	North of field	#	#	#	#

APPENDIX F

INSTRUCTOR GUIDE WITH SCENARIO PROTOCOL, TIMELINE, AND RATING CRITERIA

Scenario Instructor Guide with Scenario Protocol, Timeline, and Rating Criteria

Time Allotted:

- **Crew Brief (10 minutes)**
- **Event (20 minutes)**
- **Debrief (40-60 minutes)**

Prerequisites: NATOPS qualification and completion of the FRS Familiarization Phase.

Event Objective: Perform within acceptable standards for safety of flight and coordination for flight operations required to takeoff from NASNI to W291 and return to NASNI under normal and emergency conditions.

Evaluation: Instructors will use evaluation criteria set forth in the evaluation document attached to conduct process and outcome evaluations. One set of instructors will evaluate and score aircrew process performance and another set of instructors will evaluate and score aircrew outcome performance. Standards are set in accordance with NATOPS, SOP, and best practices.

Video and data files may be used to reconstruct and evaluate aircrew performance in the scenario
The Debrief/Interview will be recorded with a tape player.

Consent Forms/Research Brief:

- **Researcher will explain the purpose of this scenario event and its conduct prior to the event. Consent forms shall be obtained and any research-related questions answered prior to the brief.**

Trainer/Debriefing System Set-up:

- **The OFT will be conducted with the Pilot and COTAC positions. CADS is required.**
- **The aircrew will control the aircraft during the entire period.**
- **The aircrew will make all the required radio calls to the appropriate agency on the correct frequency**
- **If equipment malfunctions during the first five minutes of the scenario event the trainer will be reset and the event will be restarted from the initial conditions. If the trainer malfunctions after the first five minutes another scenario will be substituted for training purposes but will not be included in the research study.**

- ***Environmental Settings*:**
- **Night, VMC Conditions at NAS North Island**
- **Temp- 20 degrees C.**
- **Winds 290/12**
- **Landing and Departing Runways 29 and 36**

Aircraft Configuration:

- **Load out: Clean**
- **Weight: 40,000 LBS – (10,000 fuel/30,000 A/C)**
- **Crew positions occupied: Pilot and COTAC (Co-pilot)**
- **Initial position: Runway 29**

Event Brief to Crew:

Takeoff position RWY 29 with both engines running. Complete Takeoff Checklist. Takeoff on Runway 29 NZY on NASNI #2 departure to W 291. Operate in W291 for 0+30. Return to NYZ for VFR entry full stop landing.
Instruct crew to conduct Safety Briefing (approx. 5 minutes) prior to trainer event.

Directions for introducing Tower Comms/ Abnormalities:

- **Instructors will introduce Tower communications and aircraft abnormalities in accordance with scenario timeline/script. Time stamp insertion of abnormal indications provided by instructor using the Computer Aided Debriefing System (CADS) marker. Also, mark crew verbal response/action upon recognition and completed response to cue.**

Evaluation: Combine general standards, level of thought considered, and specific event set criteria.

Debrief (recorded): **instructors and/or researcher may ask Debrief questions. Please remind crews that the debrief/interview is being recorded and to project voices. CADS will be used as appropriate.**

- **The following questions will be asked in this order with clarifications requested at the instructors' /researcher's call.**

1. **How do you think you did?**
2. **What would you like to talk about first?**
3. **What were the difficult decisions for you?**
4. **Why was each decision difficult?**

. **For each decision:**

(a) What was your degree of confidence in your situational awareness (use a scale of 1-10, with 10 being the most confident)?

(b) What were the reasons for that level of confidence?

(b) How many options were considered before you choose your course of action?

© Why did you choose that particular course of action?

(d) What other actions/tradeoffs did you consider?

(e) What did you think possible explanations were for conflicting or uncertain info?

(f) What one piece of missing information would have helped you most?

(g) What would you do differently if you were in this situation again?

5. As a crew, what were your biggest strengths? What could be improved? Would you change your brief?

6. What were some important "lessons learned" from this scenario exercise?

- **Please remind crew not to discuss scenario details with other crews.**

Process Score Considerations - General Rating Scale

Scale	Rating	Definition
1	POOR	Observed performance is unsafe and potentially detrimental to the safe and orderly outcome of the event. This includes instances where necessary behavior/procedures were not present and examples of inappropriate behavior that were/could be detrimental to safety of flight and/or flight operational effectiveness.
2	BELOW STANDARD	Observed performance meets minimum requirements, but there is room for much improvement. This level of performance is less than desired for effective coordination and safety of flight and flight operations considerations.
3	STANDARD	Observed performance promotes and maintains coordination and safety of flight effectiveness. This is the level of performance that should normally occur during flight operations.
4	ABOVE STANDARD	Observed performance is significantly above expectations. This includes instances where necessary behaviors/skills were present, and demonstrated performance was instrumental in safety of flight and flight operations.
5	EXCEPTIONAL	Observed performance represents a high level of skill in the application of certain behaviors and serves as a model for coordination, teamwork, and highly efficient flight operations.

Process Score Considerations (Level of Thought Exhibited)

1. **Systems Knowledge – cue/strategy associations; aircraft safety requirements**
2. **Aircrew Coordination – Cross talk / in-flight ORM**

- **Information Exchange – Articulate problem. Utilize all available sources of info, passing info without prompts, provide periodic situation updates which summarize big picture, ask for status info/status as required**
- **Communication – Proper phraseology, completeness of reports, brevity, clarity**
- **Supporting Behavior- monitoring and correcting crew errors, discrepancies, avoiding task saturation/assisting with task shedding.**
- **Initiative/ Leadership - Take appropriate action in task prioritization/organization._Provide guidance/suggestions to crew or external agent. State clear and appropriate priorities. Work internal/external issues to alter or make plan work. Challenge assumptions.**

Process Scores-Specific Standards for each Event Set

SCENARIO- CREW #____/DATE____

Event Time	Event/Tower Comms	Standard	1	2	3	4	5	Eval	Notes
0-6 min	Takeoff/Departure Tower- *Standard Departure Comms*	All checklist items completed	1	2	3	4	5	Eval	

6-7 min	**Abnormal Indication**: ***Established at 6000ft/Heading 240*** **STARTER CAUTION LIGHT ON #1 ENGINE**	**Execution of NATOPS memory items (TFI) followed by rest of checklist**	**1**	**2**	**3**	**4**	**5**	**Eval**	
8-9 min	**Checklist Interruption** **When crew is going through No. 1 engine secure checklist:** **TOWER:** ***Be advised Viking aircraft at 30 miles with a pending emergency."***	**Complete Checklist**	**1**	**2**	**3**	**4**	**5**	**Eval**	**Notes**
9min	**Low Oil Pressure Light On No. 2 Engine** *Below 30 PSI* *If crew does not notice –surging NG*	**Recognizes low oil pressure engine only operating engine** **Reduce throttle to save engine life** **Internal//external com – need for immediate landing. Verbalized.**	**1**	**2**	**3**	**4**	**5**	**Eval**	**Notes**
10-12 min	**Approach back to NASNI** If crew asks for arrested landing: **TOWER:** *" What is your weight for arrested landing?"*	**Crew initiates/ completes the following:** **(1) Approach Checklist** **(2) SE Landing Checklist** **(3) Arrested Landing Checklist**	**1**	**2**	**3**	**4**	**5**	**Eval**	**Notes**

13 min	If crew does not recognize low oil PSI on final approach then: **TOWER:**" *Misty__ you have white smoke coming from your starboard engine.*"								
14 min	**Approach Priorities: Aircrew coordination**	**Contingency internal cockpit brief: loss of #2 Engine/waveoff, application of max power on #2 ENG** **Brief loss of #2 ENG/Waveoff/Hook skip**	**1**	**2**	**3**	**4**	**5**	**Eval**	**Notes**
19-20 min	**Foul Deck** **TOWER:** ***"Crash crew states the arresting gear is derigged with a burnt out motor. State your intentions.*** **If crew continues:** **TOWER:** ***"Crash truck is stalled at intersection of Runways 29 and 36."***	**Option 1- crew elects to wave- off.** **Full Throttle = Failed Engine**	**1**	**2**	**3**	**4**	**5**	**Eval**	Notes
		Option 2- Crew acknowledges new info. Lands anyway	**1**	**2**	**3**	**4**	**5**	**Eval**	**Notes**
		Option 3- Other action by crew	**1**	**2**	**3**	**4**	**5**	**Eval**	

RETURN TO BRIEF/DEBRIEF ROOM

Reminder: Please record times from CADS clips shown in debrief on this sheet below with reference to discussion point/ or time.

CADS CLIPS USED IN DEBRIEF

CADS File Start/Stop Times:	Topic:	Start/Stop debrief time:
CADS File Start/Stop Times:	Topic:	Start/Stop debrief time:
CADS File Start/Stop Times:	Topic:	Start/Stop debrief time:
CADS File Start/Stop Times	Topic:	Start/Stop debrief time:
CADS File Start/Stop Times	Topic:	Start/stop debrief time:

APPENDIX G

RATERS' INDIVIDUAL AND GROUP CONSENSUS OUTCOME RANKINGS

OUTCOME RATER DEMOGRAPHICS

Table G1
Independent Performance Outcome Ratings by Individual Raters

Rater	1	2	3	4
Military Rank	CDR	LCDR	LCDR	LT
S-3 Flt Hrs	2300	2900	2145	850
Designator	NFO	Pilot	NFO	NFO
Rank Order of Crews by Individual Raters				
1 (Best)	13	13	13	13
2	16	16	16	16
3	14	14	14	14
4	15	15	15	15
5	17	17	17	17
6	5	5	5	5
7	18	18	18	18
8	11	11	11	11
9	7	2	6	2
10	4	7	7	7
11	1	1	4	1
12	3	4	2	4
13	2	3	1	3
14	8	12	3	12
15	12	8	8	8
16	6	6	12	6

Note. 1 is highest crew rank for outcome score.

Table G2
Final Consensus Outcome Rank Ordering of Study Crews by Outcome Raters

Rank Order	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Crew Number	13	16	14	15	17	5	18	11	2	1	7	4	3	8	12	6

Note. 1 is highest crew rank for outcome score.

APPENDIX H

CREW RAW PROCESS SCORES (INDIVIDUAL EVENT AND SUMMED)

Crew Raw Process Scores

		Crew 1						Crew 2					
		IP		INFO				IP		INFO			
	Rater	3	1	7	5	8		1	2	7	5	8	
Event	**Description**												
1	**Takeoff**	3	3	3	3	3		3	3	3	3	3	
2	**Starter Caution** (No. 1 engine)	3	3	3	3	3		3	2	2	3	3	
3	**Interruption during checklist** (External Comm)	3	3	3	4	3		3	3	2	1	2	
4	**Low Oil 2 Recce**	3	3	2	3	2		1	1	1	1	1	
5	**Approach Priorities**	3	3	2	2	2		3	3	2	1	2	
6	**No Gear Plan**	4	3	3	3	4		1	3	3	1	1	
7	**Foul Deck**	4	3	2	2	3		2	2	2	2	3	
	Summed Scores with total	23	21	18	20	20	102	16	17	15	12	15	75
	Average Sum Score						20.4						15

Note. IP = Instructor Pilot , INFO = Instructor Naval Flight Officer (NFO)

		Crew 3					Crew 4				
		IP		INFO			IP		INFO		
	Rater	3	2	5	4		1	2	5	4	
Event	**Description**										
1	Takeoff	3	3	3	3		3	3	3	3	
2	Starter Caution (No, 1 engine)	3	3	3	4		3	2	2	2	
3	Interruption during checklist (External Comm)	3	3	3	3		3	3	1	3	
4	Low Oil 2 Recce	2	3	2	2		2	2	2	2	
5	Approach Priorities	3	3	2	2		1	1	1	2	
6	No Gear Plan	3	3	3	2		2	3	2	2	
7	Foul Deck	2	3	2	2		3	3	2	2	
	Summed Scores with total	19	21	18	18	76	17	17	13	16	63
	Average Sum Score					19					15.75

		Crew 5					Crew 6				
		IP		INFO			IP		INFO		
	Rater	3	1	5	4		3	2	7	4	
Event	**Description**										
1	Takeoff	3	3	3	3		3	3	3	3	
2	Starter Caution (No, 1 engine)	1	3	1	2		1	2	2	1	
3	Interruption during checklist (External Comm)	3	3	1	3		3	3	3	1	
4	Low Oil 2 Recce	3	3	1	2		4	4	2	2	
5	Approach Priorities	3	2	2	2		3	2	1	1	
6	No Gear Plan	3	4	3	4		3	2	3	1	
7	Foul Deck	3	3	2	4		3	3	1	2	
	Summed Scores with total	19	21	13	20	73	20	19	15	11	65
	Average Sum Score					18,25					16,25

		Crew 7					Crew 8				
		IP		INFO			IP		INFO		
	Rater	3	2	7	4		1	2	7	2	
Event	**Description**										
1	**Takeoff**	3	3	3	3		3	3	3	3	
2	**Starter Caution** (No. 1 engine)	4	3	2	2		4	4	4	4	
3	**Interruption during checklist** (External Comm)	4	3	3	4		4	3	4	4	
4	**Low Oil 2 Recce**	3	2	2	2		5	4	4	5	
5	**Approach Priorities**	4	2	2	2		2	2	2	2	
6	**No Gear Plan**	3	2	2	1		3	2	4	3	
7	**Foul Deck**	3	3	1	1		3	4	3	2	
	Summed Scores with total	24	18	15	15	72	24	22	24	23	93
	Average Sum Score					18					23.25

		Crew 9					Crew 10				
		IP		INFO			IP		INFO		
	Rater	3	1	7	6		1	3	7	8	
Event	**Description**										
1	**Takeoff**	3	3	3	3		3	3	3	3	
2	**Starter Caution** (No. 1 engine)	2	3	2	3		4	5	3	5	
3	**Interruption during checklist** (External Comm)	3	2	2	1		3	3	3	4	
4	**Low Oil 2 Recce**	4	4	4	4		4	4	4	5	
5	**Approach Priorities**	1	2	2	2		3	3	3	3	
6	**No Gear Plan**	2	4	3	2		4	4	4	4	
7	**Foul Deck**	3	4	4	4		4	4	4	4	
	Summed Scores with total	18	22	20	19	**79**	25	26	24	28	**103**
	Average Sum Score					19.75					25.75

		Crew 11					Crew 12				
		IP		INFO					INFO		
	Rater	1	3	6	7		3	2	7	6	
Event	**Description**										
1	Takeoff	3	3	3	3		3	3	3	3	
2	Starter Caution (No. 1 engine)	3	3	3	3		2	2	3	2	
3	Interruption during checklist (External Comm)	3	3	4	4		3	3	3	3	
4	Low Oil 2 Recce	4	3	4	4		2	3	3	2	
5	Approach Priorities	4	4	3	3		3	3	3	2	
6	No Gear Plan	5	3	4	4		3	2	3	2	
7	Foul Deck	3	4	4	4		3	3	2	2	
	Summed Scores with total	25	23	25	25	98	19	19	20	16	74
	Average Sum Score					24,5					18,5

		Crew 13					Crew 14				
		IP		INFO			IP		INFO		
	Rater	1	3	6	7		1	3	6	7	
Event	**Description**										
1	Takeoff	3	3	3	3		3	3	3	3	
2	Starter Caution (No. 1 engine)	3	4	3	3		3	3	3	3	
3	Interruption during checklist (External Comm)	4	3	4	3		4	3	3	3	
4	Low Oil 2 Recce	4	4	4	4		2	3	3	4	
5	Approach Priorities	4	3	4	4		3	2	3	3	
6	No Gear Plan	5	3	5	3		5	3	4	4	
7	Foul Deck	4	3	3	3		5	4	4	4	
	Summed Scores with total	27	23	26	23	99	25	21	23	24	93
	Average Sum Score					24,75					23,25

		Crew 15						Crew 16				
		IP			INFO			IP		INFO		
	Rater	2	3	1	7	6		1	2	8	4	
Event	**Description**											
1	**Takeoff**	3	3	3	3	3		3	4	3	3	
2	**Starter Caution** (No. 1 engine)	3	3	3	3	4		4	3	4	4	
3	**Interruption during checklist** (External Comm)	4	3	3	3	4		4	4	4	3	
4	**Low Oil 2 Recce**	5	5	4	4	5		5	2	4	4	
5	**Approach Priorities**	4	3	4	4	3		4	2	5	4	
6	**No Gear Plan**	3	4	4	4	4		5	4	5	4	
7	**Foul Deck**	4	4	4	4	4		4	4	4	4	
	Summed Scores with total	26	25	25	25	27	**128**	29	23	29	26	**107**
	Average Sum Score						25,6					26,75

APPENDIX I

SUMMARY TABLE OF BETWEEN SUBJECTS TEST

Summary Table of Tests of Between Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	Df	Mean Square	F	Sig.
Crew Experience	Takeoff/Departure	.00	1.00	.00	1.00	.334
	Abnormal #1	1.05	1.00	1.05	1.97	.182
	Checklists	.39	1.00	.39	1.31	.271
	Low Oil Recognition	6.25	1.00	6.25	8.61	**.011***
	Approach Priorities	4.00	1.00	4.00	12.10	**.004***
	No Gear GamePlan	4.62	1.00	4.62	10.39	**.006***
	Final Decision	5.64	1.00	5.64	24.12	**.000****
	Process Sum Score	115.29	1.00	115.29	14.32	**.002***
	Outcome Rank	132.25	1.00	132.25	8.91	**.010***
	Ejection	2.25	1.00	2.25	18.00	.100
	Fuel State	.72	1.00	.72	.16	**.001***
Error	Takeoff/Departure	.055	14	.004		
	Abnormal #1	7.469	14	.533		
	Checklists	4.159	14	.297		
	Low Oil Recognition	10.164	14	.726		
	Approach Priorities	4.629	14	.331		
	No Gear GamePlan	6.227	14	.445		
	Final Decision	3.274	14	.234		
	Process Sum Score	112.691	14	8.049		
	Outcome Rank	207.750	14	14.839		
	Ejection	1.750	14	.125		
	Fuel State	63.478	14	4.534		
Total	Takeoff/Departure	145.563	16			
	Abnormal #1	142.500	16			
	Checklists	153.390	16			
	Low Oil Recognition	170.795	16			
	Approach Priorities	117.310	16			
	No Gear GamePlan	170.240	16			
	Final Decision	162.055	16			
	Process Sum Score	7231.582	16			
	Outcome Rank	1496.000	16			
	Ejection	8.000	16			
	Fuel State	1126.960	16			
			16			
Corrected Total	Takeoff/Departure	5.859E-02	15			
	Abnormal #1	8.519	15			
	Checklists	4.550	15			
	Low Oil Recognition	16.414	15			
	Approach Priorities	8.629	15			
	No Gear GamePlan	10.849	15			
	Final Decision	8.914	15			
	Process Sum Score	227.985	15			
	Outcome Rank	340.000	15			
	Ejection	4.000	15			
	Fuel State	64.200	15			

Note. * $p < .05$ ** $p < .001$

APPENDIX J

OIL PRESSURE DATA SETS

Temporal Data for Low Oil Pressure Recognition

	CREW 1	CREW 2	CREW 3	CREW 4	CREW 5	CREW 6	CREW 7	CREW 8	CREW 9	CREW10	CREW11	CREW 12	CREW13	CREW14	CREW15	CREW16
OIL PRESSURE # 2																
Oil Press Light ON	7:00	15:05	6:57	6:23	11:12	6:15	3:15	10:26	14:21	14:37	8:55	11:31	8:35	8:41	11:02	12:39
Oil Press Light Recognition	NO	NO	11:06	11:50	13:03	6:58	NO	10:30	14:53	NO	NO	NO	9:12	NO	11:08	14:21
Engine Fluctuation Recce	NO	NO					NO			18:30	12:25	13:03		14:11		
Eng Smoke Association	11:00	21:25					13:32					NO ASSOC				
Oil Press EP Recce Delta	4:00	6:20	3:57	5:27	2:09	0:43	10:27	0:04	0:32	0:29	3:30	N/A	0:37	5:30	0:06	1:42

APPENDIX K

CREW 6 INTERVIEW TRANSCRIPT DEPICTING DECISION MAKING WITH OPPOSING MOTIVATIONS BETWEEN PILOT AND COTAC

CREW 6 FLIGHT SUMMARY OF MAJOR DECISION EVENTS

Crew 6 Interview Transcript Depicting Decision Making with Opposing Motivations Between Pilot and COTAC

Crew 6 Interview/Debrief Excerpt:

COTAC: [A difficult decision was] "an instantaneous decision when we were trying to wave-off and started losing altitude. That instantaneous decision to eject. probably at that point, there we were still at 100, maybe 150 feet probably. I didn't see the VSI. I don't know what we had on VSI...but if we were at about 100 feet."

Pilot: "That second of looking up, I wanted to see where we were in relation to the runway. Because.. if we were...I saw that we were dropping and (inaudible)".

Instructor: "Okay. Is this where you decided to wave-off?" (Instructor starts in-flight recording file at point of interest.)

Start Simulator Flight Recording

COTAC: "We're going to land on 29."

Tower: "701. Roger."

COTAC: "I'm ready to go through the Landing Checklist."

Pilot: "Speed brakes are in."

COTAC: "Locked. Speed brakes are in."

Pilot: "Fuel is at 5.5. Hook is down. Wave the gear. Okay. Three down and locked. One mile on speed."

Stop Simulator Recording. Begin Debrief/Interview excerpt

COTAC: "Yes. We waved-off right there."

Instructor: "This is where we gave you the compressor stall."

Start Simulator Flight Recording (continued)

COTAC: Okay. Three down and locked

Pilot: Roger. Set to takeoff. Indicates 5%.

COTAC: No. 2 oil pressure is at 200. We just want to get down on deck now at this point.

Pilot: OK, let me guard you. We're at 20 PSI. You're zero on the VSI. We're level

COTAC: "200 feet"

Tower: "701. This is tower. Be advised there is a crash truck stalled at the intersection of runways 18 and 29."

COTAC 701: "Get the crash truck off the runway. We're coming on in."

Pilot: "We're getting the jet pointed over towards the water. We'll eject. Eject over the water. I over shot."

COTAC: "You still have your hyd[hydraulics] pressure. Do you still have control?"

Pilot: "Negative."

COTAC: "Eject, eject, eject!"

--- (*End simulator recording. Debriefing/Interview continues.*)

Instructor: "Well, you had a zero sink rate."

Pilot: "I had to get it cleared out to the water."

Instructor: "The water? You were on deck."

COTAC: "You know, I don't know if that would have changed but my point. That's what I'm saying I thought we were close enough to the deck that even if we would have lost our engine we could land, get the brakes on even if we would have ejected. We could have ejected right before we hit the crash trucks. We had time to get that airplane slowed down so that we could either have saved lives on deck or stopped it and we could have ejected definitely later. Now, with an airplane with nobody in it going who knows where at 100 knots...so, at that point, if we were with the aircraft, we should stay with the aircraft."

(*End of Debrief/interview transcript excerpt*)

Crew 6 Flight Summary of Major Decision Events

Starter Caution Light

(1) Flight Ops – Pilot: Announced light and then commenced bold face.

Situational Awareness (SA) - COTAC: Immediately reviews No. 2 engine indications

(2) Tactical: COTAC: Declared an emergency with ATC and then requested vectors back to North Island while requesting to maintain current altitude.

(3) Flight Ops – Pilot: Instructs COTAC to break out the Pocket Checklist (PCL)

(4) Tactical – COTAC: Asks ATC to inform North Island that they will be needing the arresting gear.

Pilot observes No. 2 Engine Low Oil Pressure

(5) Flight Ops – COTAC – Tells pilot to "keep me posted" on that

(6) Flight Ops – COTAC: Refers to PCL
Discussion on procedure identification
Vocalizes that they won't be re-starting the No. 1 engine

[Instructor comment: COTAC is overbearing]

(7) Flight Ops – COTAC: Identifies need to dump fuel
Some mutual discussion on what fuel level to dump to
Crew appears to settle upon 3,000 lbs. as a workable number

(8) Flight Ops – COTAC: Talks to Base and informs them of the situation.

(9) Tactical – Pilot: Tells COTAC to ask for a turn [to give them some time?]

(10) Tactical – COTAC: Requests to delta overhead at 3,000' in order to set up for single engine landing

(11) Flight Ops – Pilot: Announces that he is maintaining his airspeed below landing gear extension speed.

(12) Tactical – COTAC: Mentions that the No. 2 Oil Pressure is fluctuating and advises the pilot to head out over the water.

(13) Flight Ops – Pilot: "Dirty Up" [landing gear and flaps extended]

(14) Flight Ops – COTAC: Secures fuel dump "because we're at 5,000 lbs."

(15) Flight Ops / Tactical – COTAC: Briefs pilot tells him what to do in case of a Hook Skip.

(16) Flight Ops / Tactical – Pilot: "We'll be keeping it on the deck."

<u>ATC informs crew that the short field gear is not available.</u>

(17) Tactical – COTAC: Asked for availability of long field gear.
Suggested to the pilot that once on deck they could "cut the engines "as if to coast into the long field gear.

(18) Tactical / Flight Ops – COTAC: drives the discussion on runway selection

(19) Flight Ops / Tactical – Pilot: Decides to head back over the water while doing the checklists

(20) Flight Ops – Pilot: Aircraft motors around a large portion of the pattern at 200' AGL while setting up for a second approach to the runway. [*Instructor comment: This is unconventional and unsafe.*]

<u>ATC informs crew that there is a crash truck stalled on the runway</u>

(21) Flight Ops – COTAC: Tells tower to get the crash truck off the runway
Tower responds by saying that the truck can not be moved.

(22) Flight Ops – COTAC: Directs pilot to wave-off

End

APPENDIX L

STRATEGIES USED BY REPRESENTATIVE HIGH, MID-RANGE AND LOW OUTCOME RANK CREWS (13,11, AND 6)

High, Mid-Range, and Low Outcome Ranking Crew Strategies

Table L1

Specific Examples of Strategies used by Aircrews under *Conditions of Uncertainty and Increasing Time Constraints* during Loss of No. 2 Engine/No Arresting Gear

Strategies Used	**High Outcome (Crew 13)**	**Mid-Range Outcome (Crew 11)**	**Low Outcome (Crew 6)**
Procedural			
SOP/Checklists	*Pilot: "Well, we're not going to make an arrestment so we don't need to go through that checklist."*	No brief for an arrested landing, or hook skip, or wave-off contingencies. Pilot appears to advance the No. 2 engine throttle close to MRT [Military Rate of Thrust – Max Throttle/100% power]	No Approach Checklist. Did not do most of arrested landing checklist. Aircraft is observed to motor around a large portion of the wave-off pattern at 200' AGL while setting up for a second approach to the runway. [This is unconventional and unsafe].
Analytical/Weighing Options			
Situational Awareness Activity	*Pilot: "Let's try to keep it on deck because we may be losing No. 2 here."*	---	*COTAC asked for availability of landing gear*
Info (cue) use	*COTAC: "Touch down and go for (the) long field gear." Pilot: "18/36" [Runway choices] COTAC: "Not available" Pilot: "Right" [Acknowledgement]*	---	COTAC drives the discussion on runway selection **(Table continues)**

Strategies Used (Analytical)	**High Outcome (Crew 13)**	**Mid-Range Outcome (Crew 11)**	**Low Outcome (Crew 6)**
Weigh/Discuss Options and/or tradeoffs to create plan	*Pilot: If we lose No. 2 we're going to have to punch out. If the EHP goes then I might have to get the jet going to the left before we eject." "Okay?" COTAC: "Okay." Pilot: "Out over the water."*	Crew looks for other runway options, finds none and decides to press Pilot: "If we can't stop we'll have to get out of the airplane."	COTAC asks for availability of long field gear. Suggested to the pilot that once on deck they could "cut the engines" [as if to coast into the long field gear.
Plan accepted?	*C- "I agree."*	*Crew decides to continue for landing.*	Pilot decides to head back over the water while doing the checklists
Contingency Plan?	*No. Committed to land. COTAC:" We can put the hook down and drag it down the runway."*	*Discussed ejection but did not complete a formalized ejection brief.*	*Did not cover all contingencies when completing Hook Skip Brief*
Attempt to gather more info?	----	----	----
Operational pace	*Expedited*	*Expedited*	*No sense of urgency*
Adaptive/ Satisficing - NOT OBSERVED/NO DATA			
Situational Awareness Activity	----	----	----
Info (cue) use	----	----	----
Attempt to gather/delay more info?	----	----	----
Accept risk- Press ahead with focused plan	----	----	----
Contingency Plan?	----	----	----
Accept situation with no attempt to act to effect probable outcome. No contingency plan	----	----	----

Note. A dashed line (----) denotes no activity observed or reported by the crew.

Table L2

<u>Specific Examples of Strategies Used by Aircrews under *Conditions of Uncertainty and Severe Time Constraints* with Truck in Intersection</u>

<u>Strategies Used</u>	**High Outcome (Crew 13)**	**Mid-Range Outcome (Crew 11)**	**Low Outcome (Crew 6)**
<u>Procedural</u>			
SOP/Checklists			
<u>Analytical/Weighing Options</u>			
Situational Awareness Activity	----	*Pilot knew stopping parameters*	----
Info (cue) use	----	----	----
Weigh/Discuss Options and/or trade-offs to create plan	*Crew made decision to land.*	Pilot considers Lindbergh Field. COTAC asks ATC for options. Asks about runway 36. COTAC negotiates landing on Runway 29 (Lindbergh not available).	----
Plan accepted?	*Both pilot and COTAC indicated their willingness to land during the debrief.*	Pilot decides to land with COTAC concurrence	----
Contingency Plan?	*Yes*	COTAC- Getting stopped in time and/or steering around the truck	----
Attempt to gather more info?	*No*	Yes. From ATC	COTAC tells tower to get the crash truck off the runway. Tower responds by saying that the truck can't be moved.
Operational pace	*Accelerated but still purposeful.*	----	----
Plan carried out	*Yes, crew proceeded with approach.*	*Crew lands. No speed brake extension.*	*Crew ejects over runway* **(Table continues)**

Analytical/Weighing Options			
Strategies Used	**High Outcome (Crew 13)**	**Mid-Range Outcome (Crew 11)**	**Low Outcome (Crew 6)**
Situational Awareness Activity	*Pilot: "Knowing where the intersection was, I was comfortable landing prior to that. Knowing what the winds were right down the runway...that helps too."[In debrief]*	----	----
Info (cue) use	----	----	----
Attempt to gather/delay more info?	----	----	----
Accept risk-Press ahead with focused plan	*COTAC: "We need to land."* *Pilot: " We'll have to land and stop it."* *Pilot: "We'll put it down on the runway and really get on the brakes."*	----	----
Contingency Plan?	*No.*	----	----
Accept situation with no attempt to act to effect probable outcome. No contingency plan	----	----	----
Plan accepted?	Yes.	----	----
Plan carried out	Crew lands.	----	----

Note. A dashed line (----) denotes no activity observed or reported by the crew.

APPENDIX M

SUMMARY MATRICES OF ANALYTICAL AND ADAPTIVE STRATEGY PROCESSES FOUND IN STUDY CREWS

Summary Tables of Adaptive Strategy Processes Used by Study Aircrews

Table M1
Range of Adaptive Strategies used Under Uncertain Conditions with *some* Time Constraints in the Decision Making Process

Info/Cue Search	Info/Cue Use	Decision Making (DM) Strategy	DM Plan Communication	Plan Action.	Operational Tempo	Info/Cue SA/Search
Active Search	Accept with reflection	Press Ahead regardless with Procedural/Tactical and/or Strategic Plan (s)	Verbalized by Pilot COTAC agree/disagree	Crew agreed upon plan	Crew Maintains existing pace	Active Search
Flatline	Rote Acceptance	Create Game Plan (Procedural, Tactical and/or Strategic)	Verbalized by COTAC Pilot agree/disagree	Pilot plan	Crew Accelerate	Flatline
Avoid	Ignore	If... then Procedural, Tactical and/or Strategic Plan	Non-verbalized by Pilot	COTAC plan	Crew Delay	Avoid
Dismiss	Reject	Forced Resolution Procedural, Tactical and/or Strategic Plan	Non verbalized by COTAC	No action	Crew Split	Dismiss

Table M2,

Range of Adaptive Strategies used for *severe* Uncertainty and *severe* Time Constraints

SA Activity Info (Cue)	Info (Cue) Use	Decision Making (DM) Process	Decision Making implementation	Decision Making Communication	Plan Modification.	Operational Tempo	Task prioritization
Active Search	Accept & act with reflection	Non-verbalized Pilot (metacognitive)	Press Ahead regardless P/T/S	Verbalized by Pilot to COTAC agree/disagree	Crew agreed upon plan	Crew Maintains Pace	Cooperative team orchestrated effort
Flatline	Rote acceptance & action	Non-verbalized COTAC (meta cog)	Create Game Plan P/T/S	Verbalized by COTAC to Pilot agree/disagree	Pilot plan COTAC agree/disagree	Crew Accelerates	Mutually exclusive agreed upon efforts
Avoid	Accept & Firewall	Pilot Only	If,, then P/T/S	Verbalized by Pilot to Tower COTAC concurrence/non	COTAC plan Pilot agree/disagree	Crew Delays	Divergent efforts (Disagreement)
Dismiss	Accept & not process	COTAC Only	Forced Resolution P/T/S	Verbalized by COTAC to Tower Pilot concurrence/non-concurrence	No new action	Crew Division Agreement	None
Not done or not recognized	Accept & reject	Mutual Crew Effort	None	None	Externally accepted	Crew Division due to disagreement	
	Reject outright	None		Accepted from external source			

Note, P/T/S denotes either or a combination of Procedural/Tactical/Strategic processes used for planning and action.

APPENDIX N

SUMMARY OF CREW CHARATERISTICS RELATED TO STUDY PERFORMANCE LEVELS

Summary of Crew Characteristics Related to Study Performance Levels

Key Attributes of Successful/Optimum Performance:

Aircraft Control/ Safety
Ability to keep the big picture and accuracy of ongoing /flight status/navigation requirements/aircraft location
Systems/procedural knowledge (willingness to abandon/change routine procedure for proper reasons)
Selectively choose strategic checklists and procedures within checklists as a safety trade-off for accelerated landing,
Knowledge of aircraft capabilities and landing environment
Establish priorities
Focus on aircraft configuration/altitude/airspeed for phase of flight and situation
Familiarity with flight/landing environment
Distance from field when commencing approach plan
Intentional deviation from lens to make field
Override/modify SOP/ Checklists for safety of flight
Aviate for ejection contingency

Workload Management
Familiarity with Checklists (reliance on memory)
Error management/Self-correction
Knowing when backup for pilot is required
Knowing when back up by COTAC is required
Ability to task shed or delay new info

Attitude/Supporting Behavior
Confidence
Willingness to be assertive
Positive reinforcement of crewmember actions
Willingness to discuss discrepancies
Willingness to hear other points of view
Willingness to disagree

Situational Awareness
Systematic screening/ analysis of info (past/current/future needs)
Recognize incompleteness of info and still act under severe time constraints
Recognize need to gather more info
Sort, filter, and prioritize info quickly/efficiently

Decision Making
Establish priorities
Willingness to troubleshoot
Resolve novel problem creatively
Circumvent problem
Comfort level with choices
Ability to assess when to override/dismiss other crewmember/external input

Communication
Continuous cross-talk to keep SA calibrated between pilot and COTAC
Adjustments to initial brief/plan/contingency plan based on changing situation
Proper weight/acceptance of external communications

Key Attributes of Poor/Sub-Optimum Performance:

Aircraft Control/Safety
Allow aircraft to exceed limits of safe ejection envelope
Lack of aircraft systems knowledge- inappropriate actions under conditions, confused reason for application, technically incorrect under condition/and or any condition
Unfamiliar with basic aviation/naval terminology (e.g., starboard)

Workload Management
Task avoidance
Careless/Hyperactive (e.g., almost shut down wrong engine)
Timing inappropriate for situation (too hasty or slow)
Verbalize but use wrong action for Checklist

Attitude/Supporting Behavior
Fatalistic
Arrogance in making routine aircraft control decisions
Desire for personal comfort overriding safety
Refusing Checklist back-up request
Undermining pilot's authority

Situational Awareness
Unawareness or denial of urgency of situation
Consumed with irrelevant issues

Decision Making
Inflexibility -Use of lens as landing aid when landing requires deviation to make runway touchdown for shortest stopping distance
Lack of intuition – land using sight of runway in windshield
Poor use of available time
No/poor strategic planning and/or strategic contingency planning (internal/external)
Limited comfort level
Mixed motivation between crewmembers

Communication
Not verbalizing issues
Use of expletives that may be misinterpreted as call for ejection
Blind acceptance of external communications/info
Inappropriate communication (e.g., "yadda, yadda, yadda", "one potato, two potato" to complete checklist)