U.S. Army Research Institute
for the Behavioral and Social Sciences

Research Report 1601

Computer Simulation Model of Cockpit Crew Coordination: A Crew-Level Error Model for the U.S. Army's Blackhawk Helicopter

William E. Griffith
Micro Analysis and Design

John E. Stewart II
U.S. Army Research Institute

September 1991

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

MICHAEL D. SHALER
COL, AR
Commanding

Research accomplished under contract for the Department of the Army

Micro Analysis and Design

Technical review by

Gabriel P. Intano
Dennis K. Leedom
Dennis C. Wightman

NOTICES

DISTRIBUTION: Primary distribution of this report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-POX, 5001 Eisenhower Ave., Alexandria, Virginia 22333-5600.

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

Research Report 1601

Computer Simulation Model of Cockpit Crew Coordination: A Crew-Level Error Model for the U.S. Army's Blackhawk Helicopter

William E. Griffith
Micro Analysis and Design

John E. Stewart II
U.S. Army Research Institute

ARI Aviation R&D Activity
Charles A. Gainer, Chief

Systems Research Laboratory
Robin L. Keesee, Director

U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel
Department of the Army

September 1991

Army Project Number
2Q162785A790

Human Performance Effectiveness and Simulation

Approved for public release; distribution is unlimited.

FOREWORD

The U.S. Army Research Institute Aviation Research and Development Activity (ARIARDA) provides human factors and training research and development support to enhance the effectiveness of Army aviators. One important application of this research is aviation safety. Two of the most safety-critical operations for Army aviators who fly air assault and utility missions are terrain flight navigation and approach and landing to confined areas under night vision goggle (NVG) conditions.

The research effort described in this report was to build a crew performance computer model using the Micro Saint modeling language. The model will predict crew navigation and obstacle clearance errors as a function of a variety of variables pertaining to crew workload, experience, mission familiarity, and coordination strategy for the UH-60 utility and MH-60K special operations versions of the Blackhawk helicopter. Micro Saint is a microcomputer-based derivative of SAINT (Systems Analysis using an Integrated Network of Tasks).

This report traces the conceptual development of the crew-level error model and presents the results of four parametric experiments performed to validate the model. All runs of the model assumed NVG conditions. Two important findings of the research were that the enhanced navigation system for the MH-60K version of the Blackhawk resulted in greatly reduced navigation error probabilities over the older UH-60 version, in spite of poor crew coordination strategies, and that crew coordination was critical to successful completion of the approach and landing phase of the mission.

This project was initiated in January 1990 by the ARIARDA Safety Team at Fort Rucker, Alabama, pursuant to the research task entitled "Reducing Army Accident Rates in Aviation and Group Operations." The completed software model was delivered to ARIARDA on 9 October 1990. Consistent with the Army's manpower and personnel integration (MANPRINT) initiative, the results of the parametric experiments with the model can generate hypotheses on how system design, crew workload, training, and crew coordination strategies jointly affect crew safety and the probability of mission success.

EDGAR M. JOHNSON
Technical Director

U.S. Army Research Institute
for the Behavioral and Social Sciences

Research Report 1601

Computer Simulation Model of Cockpit Crew Coordination: A Crew-Level Error Model for the U.S. Army's Blackhawk Helicopter

William E. Griffith
Micro Analysis and Design

John E. Stewart II
U.S. Army Research Institute

September 1991

ACKNOWLEDGMENTS

The authors acknowledge the help of several individuals who contributed to the development of the crew-level error models. We especially appreciate the assistance of the following instructor pilots (all senior Warrant Officers) at Fort Rucker: Stephen Gatewood, Phillip Garvey, John Jacobson, Ronald Littleton, Lawrence Tabert, and Michael Wheeler. They patiently explained the thought processes and procedures involved in low-level night vision goggle (NVG) navigation and obstacle clearance during approach and landing. Captain Dale Weiler was extremely helpful in describing how crews differing in mission familiarity and experience perform the terrain flight navigation phase of the mission.

Our gratitude is also extended to Carl Bierbaum and Joseph Zeller of Anacapa Sciences at Fort Rucker for providing critical task analyses for the UH-60 and MH-60K and helping to conceptualize better the linkage between pilot workload and errors.

A notice of appreciation is likewise extended to Joseph Galushka of the U.S. Army Safety Center for providing accident data from the Army Safety Management Information System database.

Several scientists from ARI Aviation R&D Activity provided invaluable assistance throughout the project. These include Jack Dohme, who donated several hours of simulator time so that we could appreciate cognitive and psychomotor task demands on helicopter pilots; Robert Wright, who provided insight into the visual acquisition of ground targets at varying distances under varying environmental conditions; and Gabriel Intano, who demonstrated the performance of NVGs.

COMPUTER SIMULATION MODEL OF COCKPIT CREW COORDINATION: A CREW-LEVEL ERROR MODEL FOR THE U.S. ARMY'S BLACKHAWK HELICOPTER

EXECUTIVE SUMMARY

Requirement:

Nap of the earth (NOE) terrain flight navigation and approach and landing to confined areas using night vision goggles (NVGs) are very difficult and hazardous operations confronting Army helicopter crews. The workload for the cockpit crew is very high, as is the consequent probability of crew-level error. Thus, one could assume that in addition to experience and proficiency, proper crew coordination and communication are essential to safe performance of these activities. This research effort sought to develop a crew performance computer model to predict navigation and obstacle clearance errors as a function of a variety of crew-level variables, including communication and coordination strategy, mission familiarity, and experience.

Procedure:

Two crew performance computer models were built for the UH-60 and MH-60K Blackhawk helicopters using Micro Saint, a microcomputer derivative of SAINT (Systems Analysis using Integrated Networks of Tasks) model language. The Task Analysis Workload (TAWL) approach was employed to model crew workload differences between the two variants of the helicopter. The missions simulated represented a helicopter flying an NVG terrain flight navigation mission that culminated in an NOE approach and landing to a clearing surrounded by vegetation of varying density. The crew-level variables (e.g., communication frequency, giving and accepting suggestions, and mission familiarity) could be set to high, medium, or low levels. These settings would in turn drive complex algorithms (based on input from subject matter experts and technical data) that would moderate the levels of intervening variables such as crew workload, timeliness of communication, and the ability to detect ground objects, thus affecting the probability that navigation and obstacle clearance errors would occur. Four parametric experiments were performed to test specific hypotheses that might have a bearing on these variables.

Findings:

The results showed that crew flying experience and familiarity (with the mission and each other) were the most important crew-level variables affecting errors related to the terrain flight navigation tasks. For the approach and landing phase of the mission, it was found that crew coordination-related variables had the strongest impact on the probabilities of obstacle clearance errors (e.g., rotor blade impacts with trees). It was also found that, in the terrain flight navigation phase, the MH-60K, with its advanced, automated navigation system, consistently outperformed the UH-60, whose crew, in spite of a Doppler navigation system, had to rely on paper maps and pilotage to navigate from one checkpoint to another. Disproportionately more UH-60s than MH-60Ks were forced to abort their missions because of navigation errors.

Utilization of Findings:

The model and its findings provide a conceptual framework for predicting and probability of success at this critical mission for crews differing in experience, mission familiarity, communication ability, and coordination strategy. The opportunity to manipulate these variables simultaneously and independently of one another seldom presents itself in most full-mission simulation and training settings. An additional benefit of these findings, applicable to full-mission simulators, is insight into the effectiveness of different crew performance measures. Consistent with the Army's manpower and personnel integration (MANPRINT) initiative, the results of the parametric experiments can generate post hoc hypotheses of how system design enhancements can overcome some of the system performance and safety limitations imposed by high workload and crew inexperience. Finally, the models for the two Blackhawk helicopter variants could serve as the foundation for a family of predictive crew-error models that could be adapted to other types of missions and helicopters.

COMPUTER SIMULATION MODEL OF COCKPIT CREW COORDINATION: A CREW-LEVEL ERROR MODEL FOR THE U.S. ARMY'S BLACKHAWK HELICOPTER

CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Computer Simulation Model of Cockpit Crew Coordination: A Crew-Level Error Model for the U.S. Army's Blackhawk Helicopter

INTRODUCTION

Overview

Purpose and scope of the crew-level error modeling effort. The purpose of this research effort was to build a crew performance computer model, using Micro Saint, a modeling framework which has been used extensively to model workload. The current model should be able to predict crew navigation and obstacle clearance errors as a function of crew workload and a variety of other variables not directly tied to workload (e.g., crew experience and familiarity, and mission threat environment). The UH-60 Blackhawk helicopter was chosen because it is a complex two-pilot aircraft. Terrain flight navigation was selected as the mission because of workload demands and the necessity that crew members communicate and coordinate with one another. This report details the conceptual development of the model and presents preliminary results from parametric experiments that were performed to validate it. Comparisons are made of workload and error probabilities for two variants of the Blackhawk: the UH-60 utility aircraft, and the MH-60K, a special operations variant soon to enter production.

Terrain flight navigation. The modeling effort discussed in this report sought to simulate a high-workload, high-risk mission profile, namely, night terrain flight navigation while evading an enemy threat, and landing in a confined area. This mission profile was modeled for both the UH-60 and MH-60K versions of the Blackhawk helicopter. Success at this mission requires close teamwork by the pilot, whose primary job is to fly the aircraft, and the copilot, whose task is to navigate by dead reckoning and pilotage, using sectional maps. When landing in a confined area, teamwork is no less important. It is incumbent on the copilot to assist the pilot in spotting obstacles that may be present in the landing zone, so that the aircraft can be cleared for landing. The use of Night Vision Goggles (NVGs) during a night terrain flight mission serves to increase an already heavy workload, making effective crew coordination even more critical to mission success. The MH-60K was chosen as a comparison aircraft because, although similar in many respects to the UH-60, it is equipped with advanced avionics and an automated navigation system which should reduce crew workload. The copilot of the MH-60K, who usually performs the navigation tasks, does not have to rely on paper maps.

That terrain flight and confined area operations impose high task demands on a mission profile for the UH-60 is borne out by a review of subject matter expert (SME) ratings of difficulty for UH-60 Aircrew Training Manual (ATM) tasks for the aircraft (Stewart & Lofaro, 1990). These tasks were found to be among the most difficult, even more so when NVGs were employed. It was also shown that operations during terrain flight and in confined areas accounted for almost a third of the pilot-error-related UH-60 accidents reported by the U.S. Army Safety Center (USASC) data base for FY 81 to FY 89.

The mission scenarios modeled were based upon simulated missions flown by the 160th Special Operations Force at Fort Campbell, Kentucky. These missions required the aircraft to follow a preplanned course, under day and NVG conditions, to fly through a hostile threat area, and to land in a confined area surrounded by trees. The simulated mission profiles had been videotaped as part of another project concerned with aircrew coordination, conducted by the Army Research Institute Aviation Research and Development Activity (ARIARDA). This provided the researchers for the present project with a useful record of instances of crew coordination, and a guide of the phases of the mission in which coordination errors were more likely to occur.

Organization and emphasis of the report. The present report details the conceptual foundation and development of the model, and examines initial results of parametric runs exercised to validate it. It explains the Micro Saint model in detail, including general model structure and operation. These details include the manner in which terrain features and mission flight are modeled. The modeling of the three dimensional aspects of approach, landing and obstacle clearance is also discussed, along with modeling of the threat environment. It discusses the various functions which express actual versus intended flight path, and visual acquisition of checkpoints by the pilot on controls. It also delineates the quantifying of timeliness and quality of required crew communications, and how deviations from expected behavior are expressed as crew navigation and obstacle clearance errors. The report also discusses the role of crew workload, mission and crew familiarity, and flying experience in the context of their impact on crew coordination and error. It delineates the assumptions made in developing algorithms used in the model. Finally, it presents the results of trial parametric runs designed to test the model's predictive capabilities.

Background

Crew level error and aviation accidents. Although the advances in aviation technology have been considerable over the past two decades, aviation accidents still occur. According to the International Civil Aviation Organization (1984), as aircraft have become safer, human error has become the dominant causal factor in commercial aviation accidents. A recent study by Boeing (1985) serves to corroborate further this finding. The Boeing research found that of all probable causes of aviation accidents over the past 10 years, fully 69% were attributable to errors and omissions by the cockpit crews.

The same trends have recently been found for U.S. Army aviation accidents. In 1990, ARIARDA examined accidents involving pilot error from October, 1983 to June, 1989 from USASC aviation accident records. It was found that the percentage of these accidents attributable to aircrew coordination errors had increased from 18% in Fiscal Year (FY) 1984 to 37% in FY 1989.

This increase may be due in part to the fact that newer Army helicopters, like the UH-60, are more complex than their predecessors and require two pilots. Another possible explanation is that nap-of-the-earth (NOE) and terrain flight have become common operational scenarios for the newer helicopters. Adding NVG-assisted night missions to this flight regime increases crew workload and makes good crew coordination even more critical to safety and mission success. Both of these proffered explanations attribute crew error to increased pilot workload which is the consequence of helicopters becoming progressively more complicated, and mission tasks more demanding.

The same ARIARDA research effort found that among helicopters requiring two pilots, the percentage of crew coordination errors is highest for the UH-60 (39%). The AH-64 was a close second (32%), and the CH-47 (16%) last. It may be of interest to note that most of the crew-level errors for the UH-60 involved errors in obstacle clearance, while most for the AH-64 concerned themselves with problems in task prioritization and workload distribution. One possible reason for the substantially lower percentage of crew-level errors for the CH-47 is the fact that this aircraft alone has a fully integrated crew consisting of pilot, copilot, flight engineer and crew chief. The crew has always been considered an integral unit for training purposes.

The criticality of good crew coordination is further underscored by preliminary findings on the impact of training interventions on the incidence of crew-level errors implicated in Navy helicopter accidents. The Navy SH-2 and SH-60 helicopter training squadrons started student training in aircrew coordination in June 1988. Alkov (in press) recently completed an analysis of Navy helicopter accidents due to crew-level errors. All observations are from July of one year to June of the next. The results of the analysis showed a drop from 51% crew-caused errors for 1987-1988, to 45% for 1988-1989, and finally to 25% for 1989-1990.

Crew coordination research. The management and communication problems confronting aircrews are similar to those which managers and team leaders face in their own work environments. Success at accomplishing team objectives depends upon the proper utilization of human as well as material resources. Thus, it is critical that communications between team members be clear and timely, that the workload be properly distributed among members, and that leaders and others in positions of responsibility be aware of situations that may affect the outcome of the collective team efforts.

Researchers such as Helmreich, Chidester & Foushee (1989) have undertaken pioneering efforts to apply these organizational and social-psychological concepts to the flight deck. Both Helmreich (1984) and Chidester and Foushee (1988) state that crew effectiveness is due to the joint effects of piloting skills, attitudes, and personality characteristics of the crew members, with those of the leader or pilot in command being the most critical. They add that until recently most researchers studying aircrews have placed the emphasis on individual piloting skills, and have tended to ignore personality and attitudinal variables. In brief, it has become obvious from the investigation of numerous aircraft accidents that even experienced and proficient crews can commit serious errors through failed communication and coordination.

Chidester and Foushee investigated the effects of two personality dimensions, expressivity and instrumentality, on the performance of aircrews in a full-mission flight simulator. Instrumentality was defined by the authors as the individual's level of motivation toward goal attainment; expressivity was conceptualized as relating to openness in communication with others. Any person can be low or high, negative or positive on either of these orthogonal dimensions. It should be noted that these two dimensions have been found to be critical to leader success across many work and team situations and are consequently part of many popular theories of leadership. Generic terms such as "task oriented" and "relationships oriented" attest to this fact.

Using a battery of personality measures, the investigators found that the samples of military and civilian pilots tested formed three distinct clusters. The first, characterized by positive traits, was high on both instrumentality and expressivity. The second was low in instrumentality but high in negative expressivity (low in achievement motivation but high in verbal aggression and complaining) whereas the third was high on negative instrumentality and low on expressivity (high in competitiveness, impatient and irritable). The validation study consisted of evaluating, in a flight simulator, the performance of 23 crews randomly assigned to captains who fit one of these three personality constellations. It was found that the best-performing crews were those whose captains were in the positive category (dubbed the "right stuff" by the authors) while the worst-performing crews were commanded by captains in the second (low instrumentality-high negative expressivity, or "no stuff") category. Crews whose captains fell into the last of the three categories (termed "the wrong stuff") showed a "sleeper effect"; their performance was poor on the first day but as good as that of the best crews on the second day. The implication of these findings is that crew familiarity can to some extent compensate for some of the problems engendered by authoritarian leadership.

Thus, in building a conceptual model of crew-level error, it is essential to go beyond the traditional notion that error occurs primarily because of excessive workload or poor piloting skills. In short, it would seem reasonable to suppose that the consequences of high workload on crew performance are moderated by individual difference variables having to do with the attitudes, personalities and communication styles of the crew members. These latter variables can directly affect the degree to which high workload situations are managed by the aircrew. A closely related issue which provided the impetus for the present research project was the question of whether a workload modeling framework, such as Micro Saint, could provide an adequate basis for predicting cockpit errors. As was previously stated, not all crew-level errors are driven solely by workload; thus, it would seem a worthwhile undertaking to determine if Micro Saint can also incorporate non-workload related individual and crew-level variables into the same framework traditionally used to predict workload. In short, workload would become an intervening variable mediated by others, such as mission and crew familiarity, and communication frequency, for the purpose of predicting crew-level errors.

METHOD

Overview

Originally, this effort called for models to predict different types of crew errors as a function of a wide variety of variables. Variables included workload, crew coordination, different levels of cognitive processing, training strategy, aircrew familiarity and experience, and different levels of technical complexity. With this objective in mind, research was initiated into how these variables related to one another and to the occurrence and prediction of crew error.

After several site visitations to Fort Rucker, Alabama, the concept and scope of the effort were refined. Its goal remained the development of models which take into account critical aspects of aircrew coordination and communication and which identify situations in which specific types of crew error occur. However, the purpose and scope were modified to focus on navigation and obstacle clearance errors and to concentrate on certain variables while not including others which were seen as too difficult to quantify or correlate with error. For example, the original concept of the models called for an examination of training strategy. The final versions of the models do not address this amorphous and hard to quantify variable.

The original model concept called for an examination of the differences in crew coordination related to tandem seating versus side-by-side seating; that is, to model both the UH-60 Blackhawk and the AH-64 Apache to compare the differences in error as a function of the seating configurations. However, after careful consideration of this issue, it was determined that there were more differences between the UH-60 and the AH-64 than simply seating configuration. These include different mission profiles, cockpit configuration, and aircraft performance parameters. Consequently, it was determined that it would be more meaningful to compare error in the UH-60 and the newer variant, the MH-60K. These two variants fly similar missions and have similar performance characteristics. The MH-60K has an advanced navigation system which is designed to reduce crew workload. It is reasonable to assume that crew coordination and communication would differentially impact performance.

At the conclusion of the concept definition phase of the model development, it was determined that the model would be built based upon the following design guidelines:

1. Micro Saint was selected as the workload modeling framework with which to construct the models.

2. The model was based upon mission segments, functions, and tasks as defined by the UH-60 and MH-60K task analyses (Bierbaum, Szabo, & Aldrich, 1989) and as modified by ARI work in progress underway at Fort Campbell, Kentucky.

3. The Task Analysis of Workload (TAWL) approach was employed to model crew workload during the UH-60 and MH-60K missions.

4. The workload-based model would predict crew navigation error and obstacle clearance error as a function of several variables, including crew coordination, mission familiarity, crew familiarity, experience, and two composite personality factors (Helmreich et al., 1989), which are termed instrumentality and expressivity.

5. The performance measures in the model would be crew navigation error and obstacle clearance error.

The remainder of this section presents a discussion of each of these topics.

The Micro Saint Model

The computer models of the mission scenario were built using Micro Saint, an MS-DOS compatible simulation software system developed by Micro Analysis and Design. Micro Saint was chosen as the simulation system principally because it was used to model crew workload for a proposed family of Army light helicopters (Laughery, Drews, Archer, & Kramme 1986). This modeling effort concerned itself with the prediction of crew workload and did not deal with the relationship of the former to crew-level error.

It is also the software "engine" for a major part of ARI's ongoing manpower and personnel integration (MANPRINT) methods project (also known as HARDMAN III), where it is used to generate workload-driven estimates of crew size, maintenance manpower, and other system performance and support requirements for new Army systems.

Stochastic vs. deterministic models. Micro Saint is based upon a stochastic as opposed to a deterministic or mathematical model. Deterministic models are typically mathematical representations of key variables that describe system performance. In these models, analysts predict system performance by assigning values to the various parameters, then solving for the value of a function. Deterministic models are usually descriptive of system performance but are hard to make prescriptive. They are typically coded in computer language and will run relatively quickly. Deterministic models are limited in that they have difficulty in accounting for complex interactions between variables. Although this limitation can be moderated by careful design and implementation, the critical limitation is that these models do not permit a dynamic assessment of system performance.

Stochastic models, in contrast, do support dynamic assessment of system characteristics and performance. They employ Monte Carlo techniques, thereby introducing an element of randomness into the interrelationship between system variables. This dynamic nature more accurately reflects the nature of complex interactions between variables.

It was necessary to represent and account for the complex interactions between crewmembers, the helicopter, and the environment in which it was flying. Hence, a stochastic simulation system was selected. Some of the dynamic interactions that were modeled included (a) competition for constrained attentional resources, (b) crew reaction to situational variables (e.g., deviations from course and the presence of threats), and (c) realistic crew coordination.

Micro Saint is a microcomputer descendant of SAINT (Systems Analysis using Integrated Networks of Tasks). It is a discrete event simulation tool that is based on the theory of task network modeling. In discrete event simulation, the system progresses sequentially from one activity or event to the next without progressing through all instances of time. This is in contrast to continuous event simulation systems which progress through time at a fixed pace regardless of whether a discernible event occurs at each instance. In essence, discrete simulation is driven by the occurrence of events while continuous simulation is driven by the clock.

Micro Saint task network. Task network modeling involves the decomposition of system performance into a series of subactivities or tasks (e.g., a task analysis of human operators, a functional analysis for a tank or an aircraft). The sequencing of tasks is defined by constructing a task network. The level of system decomposition (how finely the tasks can be defined) depends on the particular problem. The system can be defined in as detailed or gross a level as desired and common sense on the part of the modeler is usually sufficient for determining this level. A task network may include several relatively autonomous subnetworks which, while interrelated, are also distinctly separate. In the two current helicopter models, the highest level network consists of two components; terrain flight navigation and approach and landing.

In these models, once the network was defined, the tasks were "loaded" with variables which represented system behavior. For example, in these models, the tasks are loaded with information such as task performance constraints, workload values, and consequences of task performance. Additional information required to make the model "run" includes the following for each task:

1. The mean time required to perform a task.

2. The nature of the distribution in performance time (normal, gamma, rectangular).

3. The state of the system before the task can begin (crewmember availability, the location of the helicopter).

4. The task(s) which will follow when the current task is completed.

5. For tasks which may be followed by several tasks, the logic associated with selection of the task(s) which will begin after this one is completed.

6. Priority tasks which are permitted to interrupt the performance of other, lower priority tasks.

After building the task network and loading the tasks with appropriate variables, the model is ready to be executed to simulate operation of the system (including the activities of the crew within the helicopter and behavior of the helicopter within the flight environment).

The task network framework. After selecting Micro Saint as the simulation system, the next step was to determine the structure of the networks and constituent tasks which would represent the mission activities. The two options were either to select an existing mission structure or to develop one from the ground up. After a review of the available mission structures, it was determined that the UH-60 and MH-60K task analyses performed by Bierbaum, et al. (1989) and Bierbaum and Hamilton (1990), provided the best framework for the model. These analyses decompose several aviation missions into their constituent activities and concomitant time standards, workload ratings, and task performance constraints (i.e., functions and tasks which may not be performed concurrently).

Phases represent the broadest category of mission-related behavior. They correspond to the highest level mission behaviors such as departure and en route activities, and are composed of segments, which are more specific. Examples of segments include takeoff, NOE flight, and landing. Segments are themselves divided into behaviors called functions. Functions represent more specific classes of related behaviors and include such activities as checking avionics systems, monitoring flight controls, and establishing a hover. Functions are composed of lowest level of behaviors called tasks. Tasks represent individual crewmember-level activities usually performed using a single subsystem. They are typically of rather short duration and include such behaviors as activating individual controls, checking specific indicators, and transmitting or receiving communications.

The phases, segments, functions, and tasks identified by Bierbaum were easily transferred into a Micro Saint model. These included virtually all the task and function sequencing, task times, and workload ratings from the TAWL analyses. Some functions and tasks were modified after discussions with Army helicopter pilots indicated that sometimes crew behaviors differed from that specified in the Bierbaum analyses.

Once the general model framework was developed, based upon the Bierbaum analyses, it was determined that rather than simulating a generic mission, it would be preferable to model one that corresponded to ARI research underway at Fort Campbell, Kentucky. These missions were flown in the UH-60 simulators and consisted of pre-planned segments with checkpoints, terrain features, threats, and time and accuracy standards. Using these flight segments rendered a more realistic basis for the model.

The final model structure grafted certain key aspects of the Fort Campbell simulations upon the basic framework provided by the Bierbaum task analyses. The resulting model corresponded to Bierbaum's functions and tasks with their attendant times and workload ratings while representing the mission conditions of the Fort Campbell simulations. The workload rating system used by Bierbaum and his associates will be discussed in more detail later in this report. These workload values were integrated into the structure of the Micro Saint model.

<u>Model structure</u>. The UH-60 and MH-60K models are composed of two sequential segments: (a) low level NVG-aided flight along a pre-designated flight path comprising nine checkpoints, and (b) approach and landing at a confined area landing zone.

The activities of approach and landing are significantly different from those of low level NVG-aided flight and point-to point navigation. In the flight segment, the crew is primarily concerned with navigation and helicopter flight. In the approach and landing segments, the crew has successfully completed navigation and has turned its attention to maintaining obstacle clearance while descending into a confined area landing zone. In the models, the crew jointly performs the functions and tasks associated with low level NVG flight. Once the crew has visually acquired the landing zone, they transition immediately to the activities associated with approach and landing and with maintaining obstacle clearance.

Each of the two segments consists of three general types of subordinate networks. Each segment is composed of (1) pilot functions, (2) copilot functions, and (3) system functions. Within each crewmember's network are position-specific functions and tasks. As shown in Figures 1 and 2, the pilot and copilot perform their particular functions during the low level NVG flight segment. The pilot flies the helicopter and checks flight parameters while the copilot navigates and maintains an awareness of the instruments.

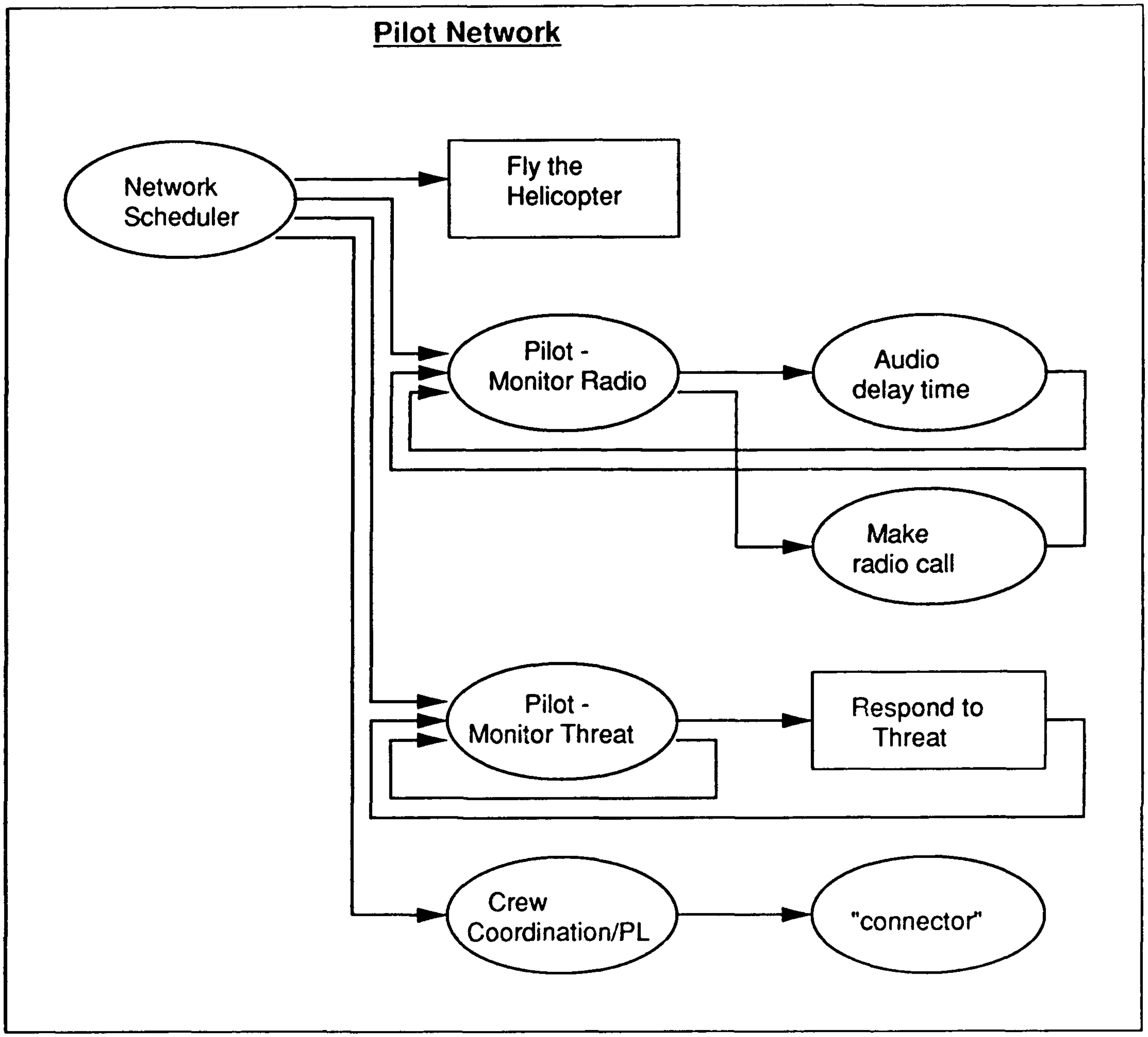

Figure 1. Pilot Network

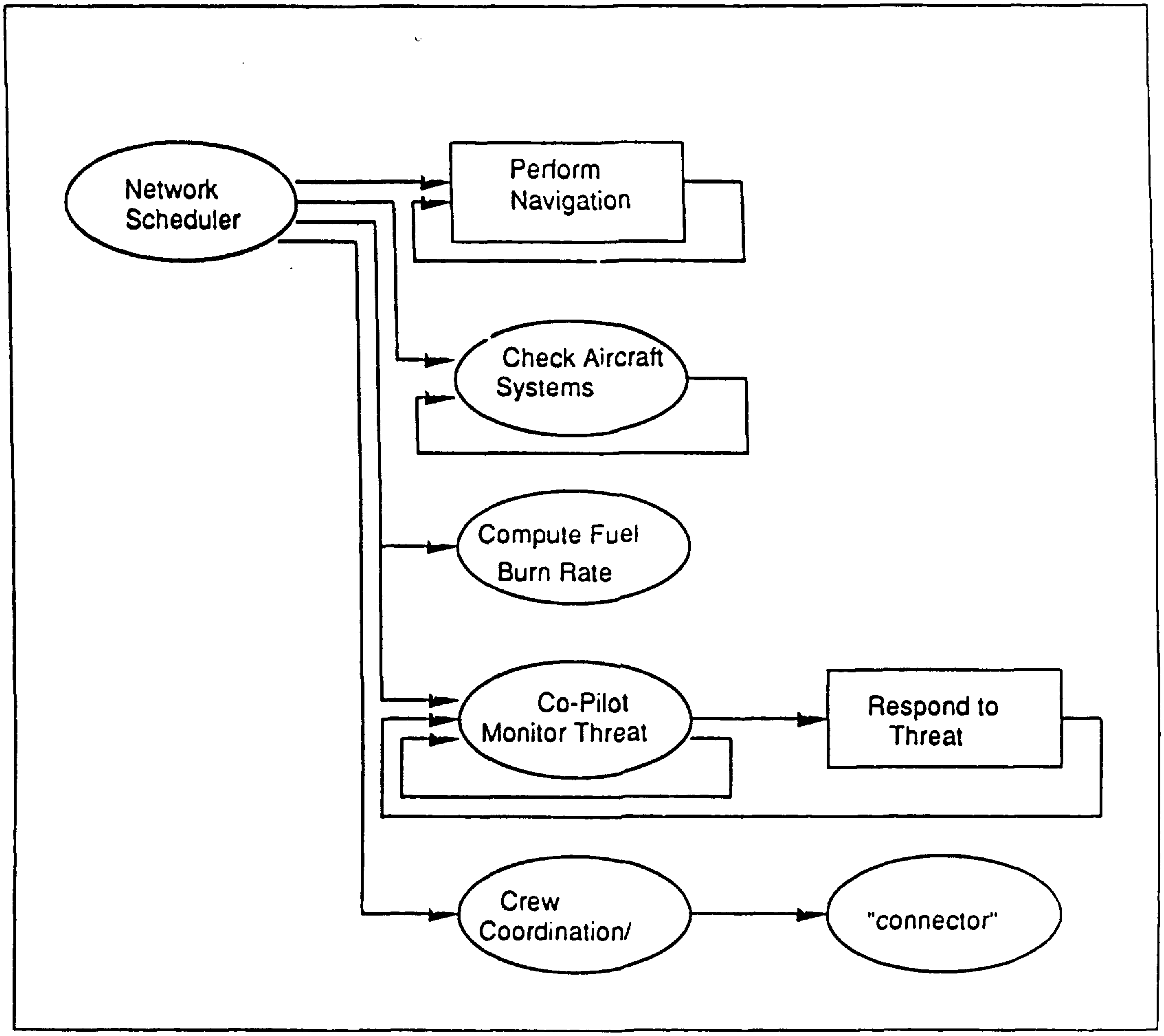

Figure 2. Copilot Network.

The system functions are activities that are performed either by the helicopter or the environment. They serve to maintain the model's "awareness" of the helicopter's position in three dimensional space and monitor the status of the helicopter relative to ground threats.

Model operations. Once the model structure was determined, it was necessary to construct functions (subroutines) and develop algorithms to cause the model to operate in a manner consistent with mission requirements. Functions and algorithms were written to perform the following operations:

1. Represent the position of the helicopter in three dimensional airspace during low level flight.

2. Determine when the helicopter deviated significantly from the intended flight path, and a means of taking corrective actions.

3. Model a threat environment that elicits appropriate defensive actions from the crew in response to threat conditions.

4. Model the approach and landing phase of the mission in three dimensional airspace where the crew must avoid obstacles while maneuvering the helicopter through trees to the landing zone.

5. Measure workload in the various channels and relate workload to the likelihood that crew coordinations will be accomplished when required, in a timely manner, and with enough informational content to serve as positive feedback to the other crewmember.

6. Quantify the timeliness and quality of both situationally-induced and routine crew coordinations and communications.

7. Relate mission familiarity, crew familiarity, and experience to the probability of timely and high-quality crew coordinations.

Representing the helicopter in three dimensional airspace. The first phase of the model simulates a helicopter flying a pre-planned route which includes a series of en route checkpoints that must be acquired visually. The model simulates a pilot flying with NVG in low level flight with his attention directed out the window to observe and report terrain features and landmarks to the copilot. The model simulates the second crewmember, the copilot, performing navigation with a map and instruments, receiving inputs from the pilot and relaying flight commands to him. The two crewmembers perform feedback functions both back to themselves and to the other crewmember, as shown in Figure 3.

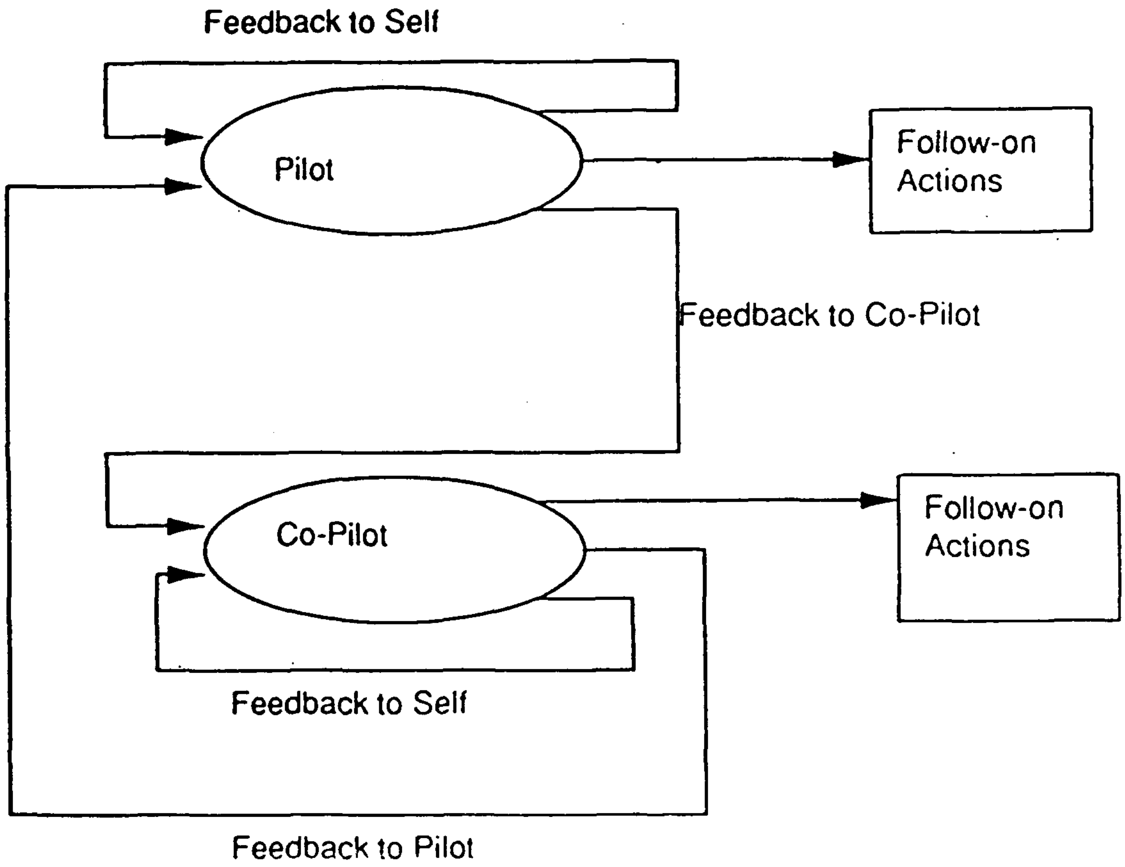

Figure 3. Crew Feedback Functions.

This portion of the model operates using several functions which monitor the helicopter's position and the pilot's ability to acquire the checkpoints visually. The helicopter position function continually updates the helicopter's position and maintains awareness of its deviation from intended flight path. It updates the position every second, taking into account airspeed, initial altitude, and azimuth and altitude fluctuations. It also calculates the deviation from the intended flight path and records as errors those instances where the deviation exceeds certain limits.

The second major function is called "DETECT" and its purpose is to simulate the pilot's ability to acquire checkpoints visually. This function derives its algorithms from the object visibility work by Wright and Debonis (1975) and takes into consideration such factors as distance and bearing to the checkpoint, angle of deflection below the horizon, tree density, airspeed, altitude, communication from the copilot, and pilot workload.

Once the checkpoint comes within range, the function evaluates all the factors to determine whether the pilot visually acquired the checkpoint on that scan. If the checkpoint is acquired, the crew plots a bearing to the next checkpoint and proceeds. Otherwise, the pilot continues to search for the point. If the checkpoint is passed without being acquired, the crew either presses on (with diminished positional certainty), executes a stationary hover to scan the area, or returns to the last known position. Each of these responses to missing a checkpoint levies a corresponding penalty on mission time or navigational accuracy.

Modeling crew responses to threat conditions. Another challenging aspect of the low-level flight segment is modeling the existence of threats and crew responses to them. The threats are represented as points from which threat sectors emanate. Within these arc segments, the helicopter is considered to be under attack. When the helicopter enters one of the threat segments, the Micro Saint model knows that it is threatened and schedules the crew to take certain actions in response. These responses involve suspending current routine flying and navigation tasks and performing evasive actions in their stead. These evasive actions involve the following changes to crew behavior:

1. Increasing the workload in several channels as a result of increased monitoring out the window and instruments.

2. Increasing the frequency with which crew communications occur.

3. Decreasing the airspeed and altitude to simulate the aircraft entering NOE flight to minimize exposure to the threat.

4. Decreasing the pilot's field of view to simulate the transition to NOE flight.

Once the helicopter has successfully evaded the threat, the Micro Saint model terminates the evasive actions and the helicopter resumes low-level flight with an increase in airspeed and altitude and decrease in some forms of workload and crew communications requirements.

Modeling approach and landing in a confined area. In the approach and landing segments, the emphasis shifts from point to point navigation to obstacle clearance functions. The model simulates the helicopter descending into a confined landing zone. The crew coordinates glide slope and obstacle clearance. The consequences of inadequate crew coordination are related to higher probabilities of obstacle strike.

During this phase, two functions are used to perform most of the critical activities. The first is called "APPROACH" and simulates the helicopter's transition from low-level flight to landing in the confined area. It calculates a glide slope and an airspeed that permit the helicopter to terminate in a hover over the landing point. This function is assisted by a subordinate function called "AVOID" which simulates the crew maneuvering the helicopter through the trees and obstacles to arrive at the landing zone having taken the path of lowest tree density. As the helicopter travels to the landing point, the model continually updates the probability of obstacle strike which is based on tree density and crew workload and coordination. In cases where tree density and workload are high and coordination is inadequate, the probability of an obstacle strike is correspondingly increased.

Representing Workload via TAWL

It should be recalled that recent task analyses by Bierbaum, et al., (1989) for the UH-60 and by Bierbaum and Hamilton (1990) for the MH-60K formed the basis for the workload model for the two aircraft. Both of these research efforts employed the TAWL approaches for assigning task workload levels to tasks performed in different segments and phases of a mission.

The initial values used for the workload ratings were based upon ratings derived from Bierbaum, et al. (1989). The authors conceptualize workload as the total attentional demand on the pilot and copilot as they perform specific mission tasks. The values are based on seven-point interval scale ratings of subjective workload estimates for each task on three component dimensions: cognitive, psychomotor, and sensory (auditory, kinesthetic and visual, both aided and unaided). Input for the actual weights assigned to each component was sought from a variety of subject matter experts (SMEs), including UH-60 instructor pilots (IPs).

The main objective of the Bierbaum, et al., (1989) effort was to develop a workload prediction model which can identify instances in which the performance of simultaneous tasks results in an overload. The level of workload is determined by summing the ratings assigned to each workload component across concurrent tasks. In an example used by the authors, imagine that a UH-60 is preparing to take off with an external load. At the 130.5 second point of the mission segment, the pilot is monitoring audio communications, interpreting hand signals from a ground guide, controlling altitude above the load, and transmitting via radio. Summing across all these tasks, it was found that the cognitive component exceeded a rating of 8.00 (the threshold for overload) with a rating of 11.0.

The TAWL workload values in the present analysis do not serve to measure workload as an end in itself, but as a mediator of crew-level error prediction. In short, instantaneous workload ratings are used to moderate the timeliness and quality of crew communications and the ability of crewmembers to perform navigation and obstacle clearance activities as a function of concomitant workload. Thus periods of increased workload will cause there to be more likelihood of crewmembers missing terrain features along their path, missing a navigational check, and will degrade the ability of one crewmember to provide proper assistance during NOE flight or approach and landing to a confined area.

Crew Coordination

Perhaps the most challenging aspect of the model was quantifying the timeliness and quality of required crew coordinations. This issue was alluded to previously in the discussion of the deviation of actual flight path from the intended path. The remainder of this section discusses the manner in which crew coordination was addressed conceptually.

Crew roles. As previously mentioned, the model simulates the crew flying a mission along pre-planned checkpoints at low level at night. The crewmembers perform different functions during this flight. The pilot is looking for terrain and cultural features to which he has been alerted by the copilot. The copilot is navigating by reading a map and informing the pilot of terrain or cultural features for which he should be watching. The pilot and copilot must communicate between themselves in order to keep on course and on time. The pilot should communicate visual acquisition of a checkpoint or failure to acquire it. The copilot should provide information to the pilot about upcoming checkpoints and features. Thus, the crewmembers serve as feedback processes to each other and to themselves. If the timeliness or quality of the communications is degraded, then each crewmember has less information to keep the aircraft on course. If crew communications suffer enough of a degradation, then the crew may deviate from the intended path significantly and become lost. These instances of significant course deviation, getting lost, and not arriving at a checkpoint within the allowed time are recorded as errors.

While it is fairly easy to describe what the model does in general terms, such as the degrading of crew coordination as a function of workload, experience, and communication style, it is quite a different matter to translate the generalities into Micro Saint expressions. One of the key activities of the second phase of this effort was to gather information about quantifying crew coordination timeliness and quality.

Modeling crew coordination. The model accounts for crew coordination by linking behaviors in several related functions into an action-feedback loop. Each crewmember performs position-specific tasks such as reading maps, following course, monitoring instruments, following ground track, and scanning for checkpoints. Each crewmember also communicates with his partner on a schedule that is determined by his individual communication style.

The model uses a complex algorithm to match performance of monitoring and situation awareness tasks with communication frequency and recency and individual workload. Depending upon the sequence of task performance, workload, and communication recency, the crew coordination can either (a) take place on time with information of sufficient quality to correct the deviation, (b) take place later than expected with high quality information transfer, (c) take place on time with degraded information transfer, (d) take place later than expected with degraded information transfer, or (e) not take place at all.

Each of the above crew coordinations may cause consequences later in the model based upon how information was passed with varying degrees of timeliness. If overload or poor crew coordination practices prevents quality information transfer, then the model simulates higher probabilities of navigation error.

Crew coordination in the approach and landing segment is different in purpose but identical in function to coordination during low-level flight. Emphasis shifts from point to point navigation to obstacle clearance functions and the model simulates a descent into a confined area landing zone. At this phase of the mission, the crew coordinates glide slope and obstacle clearance. Their coordinations take the same five forms mentioned above. The consequences of inadequate crew coordination in these segments are translated into higher probabilities of obstacle strikes (tail rotor-ground or blade-tree impact).

<u>Modeling experience, familiarity, and "stuff."</u> The model takes into account variables related to how much flying time the crew has had in the helicopter (experience) and how much time the crew has flown together (familiarity). Review of the literature showed that experienced crews can tolerate more workload demands in certain channels because they can operate more often in the schematic mode, in which complex behaviors are overlearned and

require little conscious monitoring, whereas inexperienced crews can spend more time in the attentional mode, where conscious monitoring is required (Reason, 1981).

The experience and familiarity variables are loaded into different functions. The former is loaded into the tasks, so that it moderates the effect of workload on task performance, whereas the latter is loaded into the crew coordination and communication functions. The rationale behind this distinction is that while experience is more closely related to task performance, familiarity is more closely tied to communication and interaction between crewmembers.

Besides accounting for familiarity and experience, the model also accounts for the two personality dimensions of Chidester and Foushee (1988), expressivity and instrumentality. Combined into a single variable called "stuff" (as in "right stuff", "wrong stuff", and "no stuff"), these dimensions were employed to moderate navigation and obstacle clearance performance. Because crews led by "right stuff" pilots are more open in communications and exhibit heightened motivation toward goal attainment, the navigation and obstacle clearance functions reflect higher levels of performance for these crews. (In the opinion of SMEs, obstacle clearance is mostly dependent on timely coordination and open communication between crewmembers.)

Conversely, the model penalizes crews not exhibiting "right stuff" by not moderating their navigation and obstacle clearance behaviors positively. For example if the pilot were to have the "wrong stuff", the copilot, not expecting his input to be accepted, would be less likely to communicate in a timely fashion the presence of a tree in the flight path, and the pilot, believing that a copilot's role is to be "seen and not heard" would be less likely to be attentive and responsive (or may even be indifferent) to the latter's warnings. Under this scenario, it is obvious that collision with trees and other ground obstacles is more apt to occur.

Data Requirements

Initially, it was hoped that Army aviation accident and incident data from the Army Safety Management Information System (ASMIS) would be such that we could determine the relationship between such variables as experience, crew familiarity, and mission familiarity and the probabilities of various navigational errors. This was not possible, however. While the ASMIS data were very comprehensive and complete, and USASC most forthcoming in its assistance, the data were not pertinent to the task at hand. ASMIS reports classes of aviation accidents, ranging from Class E on the low end through Class A on the high end. The performance measures required for the present modeling effort relate to navigation accuracy and obstacle clearance errors, not to their consequences.

This meant that the present investigators were not able to develop precise algorithms which related continuous values of the variables of interest (such as X pilot hours) with discrete levels of performance measures (such as Y probability of a navigation error). Instead, it was necessary to hypothesize the relationship between our performance measures and the variables of interest. Independent variables were expressed as ordinal trichotomies, (e.g., high, medium and low crew experience) with each variable ranging from a value of two (indicating a high degree of that trait or characteristic) to zero (indicating a low level). No attempt was made to equate any of the trichotomous points with any values on a continuous distribution.

This does not mean that the relationships between the independent variables and the performance measures were arbitrary. In all cases, they were the result of reviews of the literature, including a wide range of army training publications (e.g., the UH-60 Operator and Aircrew Training Manuals, numerous Flight Training Guides, and IP manuals) relevant to terrain flight navigation in the aircraft. Other sources were numerous Army aviation SMEs who were mostly standardization IPs familiar with the UH-60 and its missions, as well as interpretations of ASMIS UH-60 accident data.

Performance Measures

While many variables which measure navigation and obstacle clearance were included in the model, the focus was on three of them in this effort. Two measures were related to navigation performance: "Ne2" (navigation error type 2) and "Acq[9]" (acquisition time for checkpoint 9). Ne2 is a measure of the number of checkpoints not visually acquired during the mission, providing a measure of navigational accuracy in the point to point navigation phase. Acq[9] is a measure of time elapsed between mission start and the time that checkpoint 9, the landing zone, is visually acquired by the crew. It serves as a measure of navigational accuracy because crews who do not have to perform frequent stationary hovers or return to last known point will spend less time arriving at the landing zone.

A concrete example of an algorithm relating to navigation performance is presented for purposes of clarification. Assuming that the crew has missed a checkpoint, the algorithm is: IF cr[0,8] + cr[1,8]<=2: (cr[0,8] + cr[1,8] >2) & ((1200-time)/(10-n))>130): n==9 THEN RETURN ELSE PRESS ON. In words, this algorithm states that if crew experience is low, or if experience is high and time abundant, or if the crew is searching for the last of the nine checkpoints, then they should return to the previous known point. Otherwise (e.g., if the crew is experienced and there is not much time, and they are not looking for checkpoint 9), they should press on to the next checkpoint. This algorithm incorporates guidance received from SMEs regarding the terrain flight navigation task. Crews with low experience who do not acquire a checkpoint will come to a hover, backtrack until they visually acquire the previous checkpoint, and then head in the direction of the checkpoint that they had missed. On the other hand, experienced crews do not tend to backtrack; instead, they press on in the direction of the next checkpoint, especially when there is not much time left for completion of the mission.

The third measure is called "Pstrike" and is a measure of the instantaneous probability of an obstacle strike during the approach and landing segment. Because obstacle strikes themselves are very low frequency events, we opted to measure the average Pstrike value during the approach phase as the measure of obstacle clearance effectiveness.

<u>Rationale for Hypotheses</u>

On the basis of the foregoing rationales, several general assumptions and corollary hypotheses suggest themselves. Initial parametric runs of the model will concentrate primarily on the sensitivity of performance measures (e.g., missed checkpoints, probability of blade strike) to high and low levels of variables pertinent to crew communication and coordination, as well as differences between variants of the helicopter. The following is a list of assumptions and hypotheses that will be tested by parametric runs for the current research effort.

Crew coordination during terrain flight. Crew coordination strategy is a major determinant of crew navigation and obstacle clearance error. Examination of videotapes of UH-60 crews flying simulated missions at Fort Campbell, Kentucky suggested this hypothesis. There was remarkable correspondence between the frequency and quality of crew coordination and subsequent navigation error. Several of the subject crews engaged in infrequent, cursory crew communications. These crews invariably allowed themselves to get lost once workload increased. Several crews exhibited almost textbook perfect crew coordination. They were able to maintain an awareness of their location even during the most stressful and high workload portions of the flight.

Crew coordination during approach and landing. Based upon conversations with SMEs, it is assumed that crew coordination is more critical for the approach and landing than for the terrain flight navigation phase of the mission. They perceived approach and landing as more dependent upon good teamwork than on experience or mission familiarity. They also saw this phase as less dependent on helicopter type than the terrain flight navigation phase.

Crew experience. Another determinant of differences in navigation and obstacle clearance error is experience. Experienced crews who are also familiar with the mission should outperform inexperienced crews regardless of helicopter variant. However, it would seem reasonable to expect the effects of experience to be moderated by helicopter variant. Less experienced crews should be less likely to complete their missions successfully than would more experienced crews in the UH-60. This difference should be substantially smaller in the MH-60K than in the UH-60.

Differences due to helicopter variants. Overall, the differences in navigation errors between the MH-60K and the UH-60 should be small, but there should be differences between the good and poor crew coordinators in the two helicopter variants. While all crews may benefit to some degree from the automated navigation system (found in the MH-60K), the benefits should be most apparent in crews exhibiting poor coordination. It is these crews who should benefit most from an automated navigation system.

Crew and mission familiarity. Familiarity (both mission and crew) should more of a determinant of error rate than is experience level. The degree of familiarity has a major impact upon crew coordination. Crews that have flown together frequently have had the opportunity to develop trust and to develop and refine successful crew coordination patterns. The sense of trust is important to crewmembers. Many helicopter pilots recounted to us that they fly with more assurance with partners whom they trust, knowing that they can turn over one half of the mission to the partner without the need for constant checking. Additionally, partners who have flown together frequently develop efficient systems of communication that can compress more information into a terse response. Crew partners who are unfamiliar with one another have not developed shared expectations that make behavior predictable. Consequently, unfamiliar crew partners will be more subject to the decreased performance associated with poor crew coordination.

Hypotheses

Hypothesis 1. It would seem reasonable to expect an overall main effect of crew coordination on those dependent variables relating to navigation and obstacle clearance error. Thus, high coordination crews should miss significantly fewer checkpoints, spend less time off course, and be more likely to complete their missions than should low coordination crews. The performance advantage of the MH-60K should be greater for low than for high coordination crews. During approach and landing, high coordination crews should make fewer obstacle clearance errors when the aircraft is maneuvering through vegetation on its way to a confined landing area. Thus, the maximum probability of a blade strike should be less for high coordination than low coordination crews. Differences due to helicopter variant should be minimal during this phase. Finally, the effects of crew coordination should be stronger for the approach and landing phase than for the terrain flight navigation phase.

Hypothesis 2. This hypothesis calls for main effects of experience and (mission and crew) familiarity which should parallel that predicted for crew coordination (crews high in experience and in familiarity should make fewer errors than those low on these dimensions). For terrain flight navigation, it would seem reasonable to expect that the effects of helicopters would be greater for the low than for high experience (and familiarity) crews; consequently (Helicopter X Experience) and (Helicopter X Familiarity) interactions are predicted for those performance measures pertaining to the terrain flight navigation phase of the mission.

Hypothesis 3. This hypothesis is being proposed with substantially less confidence than the preceding two. When experience and (mission and crew) familiarity are manipulated independently of one another, it is expected that the main effect of familiarity on navigation and obstacle clearance errors will be greater than the effect of experience.

RESULTS AND DISCUSSION

Overview

The purpose of this effort is the prediction of navigation and obstacle clearance errors as a function of various design- and crew-related variables. Once the model was built, it was necessary to perform several parametric runs to determine how well the model was able to predict error. It should be emphasized that the parametric runs were executed more as a check on the sensitivity and validity of the model, not as a data collection effort for research purposes.

A total of four parametric experiments were conducted. The first experiment is a simple 20-run test of the sensitivity of three representative performance measures to different levels of independent variables pertaining to crew coordination and to helicopter type. The second experiment is a test of Hypotheses 1 and 2, and consisted of a total of 112 separate runs of the model. It comprises a full three-factor factorial design with experience, coordination strategy, and helicopter type as the independent variables. The third experiment is an attempt to manipulate experience and familiarity orthogonally as a test of Hypothesis 3. The fourth and final experiment, consisting of 64 consecutive runs, was conducted to address questions, mostly involving the effects of mission time-outs on performance, which were generated by the previous three experiments.

Experiment 1

Each model variant was executed several times, varying each time the input parameters of interest, which included (a) crew coordination frequency (20 per minute, 5 per minute, and 2 per minute), (b) expressivity and instrumentality, or "stuff" ("right stuff", "wrong stuff", and "no stuff"), (c) accepting suggestions (high, moderate, low), (d) crew familiarity (high, moderate, low), and, (e) experience level (high, moderate, low).

The performance measures collected using the model's automated data collection routines included total "Ne2" or missed checkpoints for the mission, "Acq[9]" or time to acquire the last checkpoint visually, and an average value for "Pstrike" or instantaneous probability of an obstacle strike. Micro Saint allows multiple consecutive runs; for Experiment 1, a total of 20 runs were executed in this fashion, resulting in 17 usable runs.

The comparisons deemed most important in view of time constraints on the project and the limited number of runs, were those concerning crew coordination strategy and helicopter type. One limitation on potential comparisons in Experiment 1 is the small number of completed runs (17). This was due to three of the UH-60s becoming lost. For this reason, only marginal, main-effect comparisons were made, between high vs. low crew coordination (collapsing across helicopters), and between UH-60 vs. MH-60K (collapsing across crew coordination strategies). The cell means of the previously-mentioned performance measures, along with comparisons of the marginal means via *t*-tests, are shown in Table 1 below.

Table 1

Cell and Marginal Means for Selected Performance Measures

Coordination	Variable	UH-60	MH-60K	Marginals
		Helicopter		
High	Acq[9]	830.6	907.9	878.7
	Ne2	3.2	1.0	2.1
	Pstrike	.0008	.0029	.002
	Cell n	4	5	9
Low	Acq[9]	831.5	945.9	895.1
	Ne2	3.2	1.0	2.1
	Pstrike	.012	.022	.018[c]
	Cell n	3	5	8
Marginals	Acq[9]	830.9	926.9[a]	
	Ne2	3.2	1.0[b]	
	Pstrike	.007	.013	
	Cell n[1]	7	10	

1. The reduced number of cases in the UH-60 condition is due to three of the helicopters becoming lost.

a. $p<.001$ b. $p<.0005$ c. $p<.0001$

As can be seen from the results of Experiment 1, there were significant differences between the high and low coordination groups in terms of probability of obstacle strike and between the two helicopters in terms of navigational accuracy.

In the comparison between the two coordination groups, the results are as one would expect. Those crews who exhibited a less effective communication and coordination style (i.e., fewer communications, less accepting of suggestions, "no stuff" or "wrong stuff") showed higher probability of obstacle strike than did those crews whose communication and coordination styles were more effective. No prior hypotheses were entertained concerning how Pstrike would vary between helicopter types; however, it does appear that UH-60s experienced lower probabilities of obstacle strikes during approach and landing than did MH-60Ks. In a sense, this finding is counterintuitive in that it would seem reasonable to expect that the generally lower workload for the latter helicopter should give it some advantage. In view of the small cell frequencies, interpretation of this difference should be made with caution.

The results from the helicopter comparison were somewhat mixed. In the first case, they showed that crews in the MH-60K, with its automated navigation system, were significantly less likely to commit a type 2 navigation error (missed checkpoint). This result was as expected. The other finding, however, was counter to expectations. Crews in the MH-60K seemed to take significantly longer to reach the landing zone than did those in the UH-60. This seemingly counterintuitive finding was probably due to three UH-60s becoming lost. With consecutive (as opposed to independent) model runs, dependent variables are presented as means and standard deviations. Thus missing a checkpoint results in zeros being averaged in, thus distorting the data and rendering Acq[9] virtually useless as a performance measure (it was not anticipated that this many helicopters would fail to complete their missions). Therefore, for subsequent runs of the model, separate independent runs were planned, which, though more time consuming, should yield less ambiguous data. The question of whether performance differences between the two helicopters were due to differential attrition will be addressed in Experiment 4.

Experiment 2

A second, more complex parametric experiment was conducted in order to determine the effects of helicopter type, crew coordination strategy, and mission familiarity-crew experience, on crew performance. This experiment examined crew navigation and obstacle clearance performance under conditions that are "worst case", for those crews unfamiliar with the mission and minimally experienced, and "best case", where crews were high on mission familiarity and experience.

Methodology

A total of 112 independent runs were executed, within a (2 X 2 X 2) factorial design consisting of two levels each of three factors: crew coordination strategy (high vs. low), helicopter type (UH-60 vs. MH-60K), and a composite variable of (high vs. low) mission familiarity-experience (hereinafter referred to as experience). Because of attrition due to mission failures, a total of 82 complete cases were obtained. It was expected that there would be a main effect in navigation performance due to crew coordination strategy, and to helicopter type. Crews in the high coordination condition should perform better than those in the low coordination condition, and MH-60K crews should outperform those in the UH-60. Differences in obstacle clearance during approach and landing should be due primarily to differences in crew coordination and not to helicopter type. A main effect of experience was also expected for both the navigation and obstacle clearance tasks, with crews higher in experience outperforming those lower in experience.

Variables Relating to Navigation Errors

Missed checkpoints. Table 2 below presents the results of the experiment with total missed checkpoints (Ne2) as the dependent measure.

Table 2

Summary Source Table for Analysis of Variance on Total Missed Checkpoints (Ne2) as a Function of Crew Experience, Coordination Strategy, and Helicopter Type

Source	df	MS	F
Experience (E)	1	1.60	1.30
Helicopters (H)	1	37.79	28.63b
Coordination (C)	1	.04	.03
(E X H)	1	3.48	2.89
(E X C)	1	.83	.69
(H X C)	1	5.11	4.23a
(E X H X C)	1	1.48	1.22
Error within Cell	74	1.21	

a. $p<.04$, b. $p<.01$

Table 3

Cell Means for number of Missed Checkpoints

		Coordination			
		High		Low	
		Experience			
Helicopter		High	Low	High	Low
UH-60	M	3.00	2.17	1.45	2.25
	n	11	6	11	4
MH-60K	M	.50	1.09	.86	1.91
	n	14	11	14	11

Note. Cell ns < 14 are due to mission failures.

An examination of Tables 2 and 3 shows only one significant main effect, due to helicopter variant. The MH-60K clearly outperformed the UH-60 in the terrain navigation task. This is hardly surprising, when one considers the state-of-the-art navigation system of the newer aircraft. Because there was a significant (Helicopter X Coordination) interaction, independent *t*-tests were computed for the simple effects of helicopters within each coordination level. Recall that Hypothesis 1 called for greater differences between helicopter types for poorly coordinated crews, than for well-coordinated crews. For the high coordination MH-60K ($M = .76$) and UH-60 crews ($M = 2.71$), this difference was highly significant ($t_{(40)} = -3.97$. $p < .01$). For low coordination MH-60K ($M = 1.12$) and UH-60 ($M = 2.00$) crews ($t_{(38)} = -2.32$ $p < .05$). Both were in the direction expected. However, contrary to expectations, the difference was larger for well coordinated crews. Differences favored the newer version of the Blackhawk, with even the low coordination MH-60K crews showing superior performance to the high coordination UH-60 crews. It seems that both high and low coordination crews benefited from the MH-60K's reduced workload.

Mission failures. The preceding analyses consisted of only those aircraft that were able to complete an entire mission. A total of 30 did not. Thus another variable relevant to crew navigation performance was the number of mission failures (following recommendations from SMEs, crews were timed out of the mission if they failed to acquire checkpoint 9 by 1200 sec, or were deemed lost if they missed three consecutive checkpoints while staying more than 1 km off course). The differences in failed missions seemed due to an interaction between helicopter type and experience level, with a full 24 (43%) of UH-60 missions ending in failure, as opposed to only 6 (11%) of those for the MH-60K. All of the 6 MH-60K mission failures occurred where crews were inexperienced, as did 18 (75%) of those for the UH-60. Two independent Chisquare tests of association (corrected for discontinuity) comparing mission success vs. failure rate with helicopter type were significant (X^2= 4.67, df=1, p <.05) for the high experience condition, and highly significant (X^2= 8.23, df=1, p<.002) for the low experience condition.

The implications of these findings are twofold. First, crews who are not experienced with their aircraft and mission will be much less likely to complete the mission than would those who are experienced, and secondly, those crews flying MH-60Ks, even when inexperienced, have a much better chance of completing an NVG terrain flight navigation mission than those flying UH-60s. This set of findings seemed supportive of Hypothesis 2.

Percent time off course. Another navigation variable, percent of time that the helicopter is off course (percent navigation error type 1), yielded no significant main effects or interactions due to experience, helicopter type, or crew coordination strategy.

Time to acquire final checkpoint (Acq[9]). This variable showed two significant main effects and a significant interaction. These are presented below in Table 4. A glance at Table 4 reveals a very large main effect of experience, with experienced crews arriving at the final checkpoint sooner (M= 857.48 sec) than inexperienced crews (M= 1011.34). The main effect of helicopter type was marginally significant, with MH-60Ks arriving earlier (M= 912.00) than UH-60s (M =926.16). The only significant interaction was (Experience X Helicopter), with mean differences due to helicopters being much greater for low experience(Ms= 988.32, MH-60K; 1062.00, UH-60) than for high experience (Ms= 852.04, MH-60K; 864.41, UH-60). These simple effects were compared via independent t tests. The differences in helicopters under low experience approached significance for 38 degrees of freedom (t= -1.83, $p < .07$), and the simple effects of helicopters under high experience were not significant for 40 df (t= -1.41, $p<.14$).

Table 4

Summary Source Table for Analysis of Variance on Time to Acquire Final Checkpoint (Acq[9]) as a Function of Crew Experience, Coordination Strategy, and Helicopter Type

Source	df	MS	F
Experience (E)	1	484891	107.72c
Helicopters (H)	1	23653	5.23b
Coordination (C)	1	10013	2.22
(E X H)	1	18994	4.22a
(E X C)	1	7283	1.62
(H X C)	1	18	<1.00
(E X H X C)	1	453	<1.00
Error within Cell	74	4501	

a. $p<.04$
b. $p<.025$
c. $p<.001$

Summary of results pertaining to navigation errors. It appears that the results with regard to the effects of crew coordination on navigation error did not support Hypothesis 1, which called for a main effect of crew coordination level on navigation error, but weakly supported the second corollary of the hypothesis that called for an interaction between crew coordination strategy and helicopter type. Among highly coordinated crews, it was expected that overall differences in navigation performance due to helicopter type would be minimal; however, among poorly coordinated crews, it seemed reasonable to expect that MH-60K crews would outperform UH-60 crews. Navigation errors between the two types of helicopters were anything but minimal, with the MH-60K showing a clear advantage over its predecessor.

All in all, it seemed that experience was a stronger factor than crew coordination strategy. Only one main effect of coordination approached significance, and there was only one significant interaction (with helicopter variant) and in this the simple effects of helicopter variant favored the MH-60K, regardless of crew coordination level.

Variables Relating to Obstacle Clearance

A corollary of Hypothesis 1 predicted a main effect of level of crew coordination on landing phase errors involving obstacle clearance while the helicopter is maneuvering through vegetation varying in density (blade strikes, probability of blade strike, and maximum probability of blade strike during the approach and landing phase). It was also expected that the main effects of coordination would be larger than those of the other two independent variables. No main effect due to helicopter type is predicted.

Maximum probability of a rotor blade strike. A dependent variable closely related to Pstrike is the maximum probability of a blade strike, or Psmax. This represents the high water mark during the approach and landing phase of a mission. A three factor analysis of variance, comprising the same factors as before, was performed on this variable. The main effects and interactions are presented in Table 5 below.

Table 5

Summary Source Table for Analysis of Variance on Maximum Probability of Rotor Blade Strike (Psmax) as a Function of Crew Experience, Coordination Strategy, and Helicopter Type

Source	df	MS	F
Experience (E)	1	.003	6.62b
Helicopters (H)	1	.006	11.98c
Coordination (C)	1	.046	94.49c
(E X H)	1	.001	1.77
(E X C)	1	.001	1.69
(H X C)	1	.016	32.40c
(E X H X C)	1	.003	4.91a
Error within Cell	74	.0005	

a. *p* <.03, b. *p*<.01, c. *p*<.001

Unlike the terrain flight navigation errors, the strongest main effect for Psmax was due to crew coordination strategy. This effect was in the expected direction, (*M*s= .01, high; .08,low), and supportive of Hypothesis 1. There was also a significant (Helicopter X Coordination) interaction and a complex, uninterpretable (Experience X Helicopter X Coordination) interaction.

Differences between helicopters were not predicted for the approach and landing phase of the mission. Thus, the (Helicopters X Coordination) was unexpected. The simple effects of helicopters were probed at each level of coordination via independent *t* tests. For high coordination, the difference between the MH-60K (*M*=.008) and UH-60 (*M*= .017) coordination conditions was nonsignificant ($t_{(40)}$= <1.00). For low coordination the same comparison between the MH-60K (*M*=.078) and UH-60 (*M*=.029) conditions yielded a *t* ratio of -8.30, which for 38 degrees of freedom, was highly significant (*p*< .0001). It would seem, then, that the obtained interaction disconfirmed the expectation that there would be no differences between helicopters for this particular phase of the mission.

Although in general high coordination crews outperformed low coordination crews, low coordination crews were quite different with regard to helicopter type, with UH-60 crews outperforming those in the MH-60K. For high coordination crews, helicopter type seemed to make no difference. This simple effect was consistent with predictions.

The significant main effect of helicopter types ($\underline{M}$s=.04, MH-60K; .02, UH-60) is interesting, though unexpected. There was no prior rationale for expecting that the UH-60 would show a lower maximum probability of blade strike during approach and landing than would the MH-60K. A possible explanation is that most poorly performing UH-60 crews were culled out by failing to complete the mission, whereas most poorly performing MH-60K crews were able to survive the entire mission. In short, it could be that the advanced navigation aids of the MH-60K gave poorly performing crews a definite advantage over their UH-60 counterparts; however, this advantage disappeared during the approach and landing phase, when they could no longer rely on high technology to assist with this latter phase of the mission.

Finally, it should be mentioned that the significant main effect of experience was in the direction expected ($\underline{M}$s= .03, high; .04, low), supporting Hypothesis 2.

<u>Actual number of rotor blade strikes</u>. In addition to computing probabilities of blade impacts, the model also simply counted the number of actual blade strikes during the approach and landing phase of a mission. Table 6 presents the results of the ANOVA on this particular variable. As can be seen from an examination of the table, the only significant main effect was due to coordination level.

There were two significant interaction effects: the first, a marginally significant (Experience X Helicopter) interaction, and the second, a highly significant (Coordination X Helicopter) interaction. These were considered of interest, since no differences in obstacle clearance due to helicopters were anticipated. Mean differences for the coordination main effect were in the expected direction with the 42 high coordination crews experiencing an average of .31 strikes per mission, versus 1.75 for the 40 low coordination crews.

Table 6

Summary Source Table for Analysis of Variance on Number of Rotor Blade Strikes as a Function of Crew Experience, Coordination Strategy, and Helicopter Type

Source	df	MS	F
Experience (E)	1	.375	3.35
Helicopters (H)	1	.369	3.31
Coordination (C)	1	5.020	44.92b
(E X H)	1	.458	4.10a
(E X C)	1	.175	1.56
(H X C)	1	1.700	15.21b
(E X H X C)	1	.113	1.02
Error within cell	74	.112	

a. p <.05, b. p <.001

Because there was a significant (Helicopter X Coordination) interaction, the simple effects of helicopter type were probed for each coordination level via independent t-tests. For the high coordination crews, mean differences between MH-60K (.59) and UH-60 (.12) were not significant ($t < 1.00$). For the low coordination condition, the respective means (2.32, MH-60K and .80, UH-60) were significantly different ($t_{(38)} = 3.88$, $p < .025$,) in the direction of fewer blade strikes for the UH-60.

Similarly, the simple effects of helicopters were examined under each level of experience. For low experience crews, means were not significantly different (1.18, MH-60K; 1.40, UH-60);($t_{(30)} < 1.00$). For high experience, the difference due to helicopters was marginally significant ($t_{(48)} = 2.42$, $p < .05$) with the mean for the MH-60K (1.25) being higher than for the UH-60

(.36). It seems, then, that this interaction was due to the unusually low cell mean for the UH-60 in the high experience condition. This further supports the post hoc explanation of mission attrition varying systematically with helicopter type.

Summary of results pertaining to obstacle clearance errors during approach and landing. It appears that Pstrike in Experiment 1 and Psmax and number of blade strikes in Experiment 2 yielded very similar effects due to differences in crew coordination strategy. These findings confirmed Hypothesis 1. Across Experiments 1 and 2, the evidence is consistent with regard to the crew coordination factor, but inconsistent with regard to helicopter type, no differences having been expected. The inconsistency could in part be explained by the smaller number of UH-60 observations in both experiments, due principally to the lesser likelihood of the UH-60 to complete a mission. The relatively higher mission "survival" rate of the MH-60Ks could account for the overall higher probabilities of blade impact for this aircraft during the approach and landing phase of the mission. It could very well be that UH-60s which would have shown worse performance, never got a chance to initiate this phase of the mission.

Correlations between Variables

Table 7 presents the intercorrelations between the three independent variables and six performance measures which comprised the dependent variables for Experiment 2. For the entire set of variables, *df*= 80. A glance at Table 7 shows patterns of correlation that are consistent with the findings in Experiment 3. It appears that Experience is most highly correlated with those variables related to the terrain flight navigation task, whereas Coordination is highly correlated with variables pertaining to obstacle clearance. Helicopter type is most highly correlated with the number of missed checkpoints.

As one would expect, the correlations between the three independent variables should be close to zero. However, note that the correlation between Experience and Helicopters, appears to approach significance (r=-.13). This implies that the more experienced the crew, the more likely it is to be flying the UH-60. This was probably due to the differential mission attrition rates for UH-60 versus MH-60K crews, and will be tested in Experiment 4.

Table 7

Pearson Intercorrelations of Experience (E), Helicopters (H), and Crew Coordination (C) with Mission Time (T), Missed Checkpoints (MC), Percent Time Off Course (OC), Acquisition Time for Checkpoint 9 (Acq[9]), Maximum Probability of Blade Strike (PS), and Number of Strikes (S)

	Variables								
	E	H	C	T	MC	OC	Acq[9]	PS	S
E		-.13	-.03	-.72c	-.04	-.21a	-.74c	-.18a	-.13
H			-.03	-.06	-.53c	-.04	-.07	.27b	.18a
C				-.09	.04	-.03	-.08	-.64c	-.50c
T					.36	.21a	.98c	.19a	.20a
MC						.13	.36b	.12	.07
OC							.19a	.05	.02
PS									.73c

a. $p < .05$
b. $p < .01$
c. $p < .001$

Experiment 3

This parametric experiment manipulated independently the variables of experience and (mission and crew) familiarity. This was a test of Hypothesis 3, which predicted that familiarity would have a greater effect on aircrew errors than would crew experience.

A total of 42 runs were executed, in a (2 X 2) factorial design, comprising two levels of mission and crew familiarity (high vs. low) and two levels of crew experience (high vs. low). Dependent variables were the time required to acquire checkpoint 9 (Acq[9]), total missed checkpoints, or Ne2, and maximum probability of a rotor blade strike. In this experiment, the variant of helicopter (MH-60K) was held constant. All of the crew-level independent variables related to crew coordination (excepting crew familiarity) were set at the high level across all conditions.

The results for the 40 complete cases presented below in Table 8 indicate no support for Hypothesis 3. It seems that experience is a stronger determinant of aircrew error than is (mission and crew) familiarity, at least for highly coordinated crews flying the MH-60K. For Acq[9], the main effect of experience was highly significant, with no other main effects or interactions even approaching significance. Experienced crews acquired the final checkpoint sooner (M= 848.26, s= 14.10, n=19) than did inexperienced crews, (M= 920.31, s= 94.30, n=19) regardless of the level of crew and mission familiarity.

For the Ne2 (missed checkpoint) variable, the effect of familiarity, though in the expected direction, only approached significance ($F_{(1,36)}$= 2.12, p<.14), with familiar crews missing fewer checkpoints (M=. 65, s=.85) than unfamiliar crews (M= 1.25, s=1.58).

All other main effects and interactions, for this and for the remaining dependent variable (Psmax), were less than 1.00.

Table 8

Summary Source Table for Analysis of Variance for the Effects of Crew Familiarity and Experience on Acquisition Time for Final Checkpoint (Acq[9])

Source	df	MS	F
Familiarity (F)	1	3104	<1.00
Experience (E)	1	49311	9.90*
(F X E)	1	322	<1.00
Error within Cell	34	4981	

*p <.01

These findings can be interpreted to mean either that it is crew experience, not familiarity that determines crew performance, or that this is true only in the case of the MH-60K. Indeed, it would be consistent with the findings of the previous parametric experiments to assume that the improved technology of the MH-60K's navigation system is "forgiving" to the crews who are less familiar with the mission and with each other. This is exemplified by the small number of MH-60Ks (n=2) which either got lost or were timed out of the mission before acquiring the final checkpoint. It would seem worthwhile, then, to examine the effects of familiarity for both helicopter variants, with the mission time out mode disabled.

Experiment 4

This final parametric experiment comprised a (2 X 2 X 2) factorial which varied (mission and crew) familiarity (high and low), a composite factor of crew coordination and experience (high and low) and helicopter variant (UH-60 and MH-60K). The coordination and experience variable will be referred to as coordination for the remainder of the experimental analysis; the reader should be aware that this variable is somewhat different from coordination as conceptualized in the previous experiments in that crew familiarity is not a component of it, but is instead manipulated independently. Crew experience is aliased with the coordination variable in this particular experiment. High coordination crews are high in experience and vice versa.

The primary way in which Experiment 4 differed from the others was in that crews which did not arrive at Checkpoint 9 in a timely fashion were not timed out of the mission, and crews who missed three consecutive checkpoints while straying over 1 km off course were not classed as lost and removed from the simulation. Thus, the time constraint on performance is removed, and even crews who fail the mission are allowed to continue their roles in the simulation.

The rationale behind this experiment was to test two post hoc hypotheses that: (a) differences in obstacle clearance performance, which clearly favored the UH-60, and (b) earlier arrivals of UH-60s at the final checkpoint (9) as found in Experiment 1, were both due to a confounding of differential attrition with helicopter variant. That is to say, the UH-60 crews who did not complete their missions were those most apt to arrive late at the final checkpoint and to perform poorly during the approach and landing phase. Since these crews were eliminated from the mission before they could execute an approach and landing, it remains moot as to how they could have performed on the latter task. MH-60K crews, however, were not selected out in this systematic fashion.

It would seem reasonable to suppose, then, that overall there will be no difference between helicopters for the obstacle clearance task (with regard to Psmax), if self-selection is not allowed to operate. As a corollary, one should anticipate a high negative correlation between the dichotomous mission pass-fail variable and Psmax. Although obvious, one should expect large differences favoring the MH-60K with respect to the Acq[9] variable. In keeping with previous predictions, it is expected that differences in terrain navigation but not obstacle clearance performance (between helicopter types) should be greater for unfamiliar than for familiar crews.

Correlations

Pearson correlations were computed between the independent variables of familiarity, coordination, and helicopter type, and five dependent variables: missed checkpoints (Ne2), Psmax, number of rotor blade strikes, time to acquire checkpoint 9 (Acq[9]), and a dichotomous variable of mission pass or fail, which indicates whether or not the crew would have been timed out of the mission or deemed lost, if this rule were enforced. Table 9, which follows, presents these intercorrelations.

Table 9

Intercorrelation Matrix for Familiarity (F), Coordination (C), Helicopters (H), Missed Checkpoints (Ne2), Maximum Probability of Blade Strike (PS), Number of Strikes (S), Mission Success or Failure (SF), and Time to Acquire Final Checkpoint (Acq[9])

Variables

	F	C	H	Ne2	PS	S	SF	Acq[9]
Familiarity	.00	.00	-.22a	-.14	-.30b	.50c	-.73c	
Coordination		.00	.04	-.75c	-.67c	.04	-.06	
Helicopters			-.05	-.11	.02	.50c	-.35b	
Missed Checkpoints				-.02	-.03	-.60c	.61c	
Max. Prob. Strike					.79c	-.06	.18	
Blade Strikes						-.07	.23a	
Mission S F							.87c	

a. p <.05, b. p <.02, c. p <.001

Table 9 shows that, as expected, allowing all crews to complete the mission, regardless of their arrival time at Checkpoint 9, resulted in orthogonality among the three independent variables. There was also a marginally significant correlation between Acq[9] and total number of blade strikes, which implies that those crews who acquired the final checkpoint most punctually were also more likely than their tardy counterparts to do well in obstacle clearance on approach and landing. It is also noteworthy that as expected, there was no significant correlation between helicopter type and Psmax or number of blade strikes during this phase of the mission. However, the correlation between the dichotomous (success-failure) variable and Psmax, contrary to expectations, was close to zero.

ANOVAS on Dependent Variables

A total of 64 runs of the model were conducted (eight conditions with eight runs per cell). The results of the ANOVAs will first be presented for dependent variables relating to obstacle clearance, and second, for those dependent variables related to terrain flight navigation. This order of presentation was chosen because the issues that prompted Experiment 4 were concerned mostly with the approach and landing phase of the mission.

Approach and Landing

The results of the ANOVA on Psmax appear below in Table 10. Recall that a main effect of coordination was predicted for this variable, with high coordination aircrews encountering significantly lower blade strike probabilities than low coordination aircrews. It should also be noted that there were no predicted differences between helicopters with regard to this variable.

Table 10

Summary Source Table for Analysis of Variance for the Effects of Familiarity, Coordination, and Helicopter Type on Maximum Probability of Rotor Blade Strike (Psmax)

Source	df	MS	F
Familiarity (F)	1	.002	3.32
Coordination (C)	1	.044	90.68c
Helicopters (H)	1	.001	1.90
(F X C)	1	.001	2.84
(F X H)	1	.005	10.00b
(C X H)	1	.001	1.84
(F X C X H)	1	.003	5.61a
Error within Cell	56	.0001	

a. *p* <.025, b. *p* <.003, c. *p* <.001

Table 10 shows a large main effect of coordination (Ms=.06, low; .00, high), but no significant main effect of helicopters. There were, however, a significant (Familiarity X Helicopters) interaction, and a highly complex, uninterpretable (Familiarity X Coordination X Helicopters) interaction. A closer look at the simple effects of the two-way interaction may prove enlightening. For obstacle clearance performance, no differences due to helicopter type were expected. The Ms for the MH-60K (.041) and UH-60 (.031) were not significantly different under low familiarity (t < 1.00). Under high familiarity these same respective means (.013, MH-60K; .039, UH-60) were significantly different ($t_{(30)}$= -2.64, p <.025).

For total number of blade strikes, the results were similar, with large main effects of both familiarity (Ms= 1.63, low; .75, high) and coordination (Ms= 2.16, low; .22, high). These results of the ANOVA on this variable appear in Table 11 below.

Table 11

Summary Source Table for Analysis of Variance for the Effects of Familiarity, Coordination, and Helicopter Type on Number of Rotor Blade Strikes

Source	df	MS	F
Familiarity (F)	1	1.13	15.02b
Coordination (C)	1	6.47	86.34b
Helicopters (H)	1	0.00	0.00
(F X C)	1	0.56	7.51a
(F X H)	1	0.52	6.90a
(C X H)	1	0.01	.14
(F X C X H)	1	0.17	2.22
Error within Cell	56	.075	

a. p <.01, b. p <.001

There was a significant (Familiarity X Coordination) interaction. Among unfamiliar crews, those low in coordination had significantly more blade strikes (M= 2.94) than those high in coordination (M= .31). For 30 df, this difference was highly significant ($t = -7.67$, $p < .005$). For those crews high in familiarity, the trend was similar (M= 1.38, low coordination; M= .13, high coordination), but only approached significance ($t = 1.60$). These differences appear consistent with the rationale calling for greater importance of coordination than familiarity for obstacle clearance tasks.

The second interaction, (Familiarity X Helicopters), was not predicted. Because there was no main effect due to helicopters, the simple effects of familiarity were probed via independent t-tests for each helicopter type. For the MH-60K, high (M=.50) and low (M= 1.94) familiarity crews were significantly different ($t_{(30)} = -2.72$, $p < .01$). But for the UH-60, familiar (M= 1.00) and unfamiliar (M= 1.31) crews were not significantly different ($t_{(30)} < 1.00$).

Terrain Flight Navigation

One of the major variables pertinent to the terrain flight navigation task is the number of missed checkpoints or navigation error type 2 (Ne2). The ANOVA performed on this variable (see Table 12) showed only one significant main effect, that of familiarity, with familiar (M= 2.09) crews missing fewer checkpoints than those who were unfamiliar (M= 2.84). There were two significant interactions. One of these (Familiarity X Helicopters), was highly significant.

Table 12

Summary Source Table for Analysis of Variance for the Effects of Familiarity, Coordination, and Helicopter Type on Number of Missed Checkpoints (Ne2)

Source	df	MS	F
Familiarity (F)	1	9.00	7.43b
Coordination (C)	1	0.25	0.21
Helicopters (H)	1	0.56	0.47
(F X C)	1	5.06	4.19a
(F X H)	1	100.00	82.66c
(C X H)	1	2.25	1.86
(F X C X H)	1	3.06	2.53
Error within Cell	56	1.21	

a. p <.05, b. p <.01, c. p <.001

An examination of the simple effects of familiarity under low coordination showed only that familiar crews (M= 1.75) missed significantly fewer checkpoints than did unfamiliar crews (M= 3.06); $t_{(30)}=-2.53$, $p < .025$). The same comparison of familiar (M= 2.44) and unfamiliar (M= 2.65) crews under high coordination yielded a t-ratio less than 1.00. Thus it seems from this experiment that the effects of (mission and crew) familiarity were in the expected direction, but only under low levels of coordination.

The size of the (Familiarity X Helicopters) interaction, which has a direct bearing on Hypothesis 2, was substantial. Recall that unlike the previous experiments, all helicopters were allowed to complete their missions, no matter how many checkpoints they had missed. It is also important to note that an interactive hypothesis was proposed for the effects of familiarity on those performance measures relating to terrain navigation. It was predicted that crews who are more familiar with the mission and with each other should be more likely to proceed on to the next checkpoint, instead of backtracking to the previously-missed one. For the MH-60K, it appears that familiar crews (M= 3.25) missed more checkpoints than did unfamiliar crews (M= 1.50); however, for the UH-60, the direction was the opposite (Ms=.94, familiar; 4.19, unfamiliar).

For MH-60K crews, this difference was highly significant $t_{(30)}$= 4.07, p <.001), as it was for the UH-60 crews ($t_{(30)}$= -8.68, p <.001).

The large interaction between familiarity and helicopters could have been due to different strategies on the part of familiar versus unfamiliar crews. Recall that the workload for terrain navigation in the MH-60K is greatly reduced. It would seem reasonable that for this helicopter, the correlation within cell under low familiarity should indicate a positive relationship between missed checkpoints and time to acquire the final checkpoint. Under high familiarity, the correlation should be close to zero.

An internal analysis of the correlation between these variables supported this post hoc hypothesis; for the MH-60K low familiarity condition the r was .89, while for the high familiarity condition it was .07. For <u>all</u> UH-60 runs, r=.93 (df=30, p <.001). Interestingly, r for <u>all</u> MH-60K runs was -.34 (df= 30, p<.06), suggesting that for this aircraft, missing more checkpoints can for familiar crews be indicative of superior performance. Across both helicopters, the overall correlation between missed checkpoints and time to arrive at the final checkpoint was .61 (df= 62, p <.001). Taken together, these findings indicate that differences in navigation strategy between helicopter types accounted for this complex interaction.

Recall that there were additional findings, which initially seemed counterintuitive, with regard to the time it took the crew to acquire the final checkpoint, or Acq[9]. In Experiment 1 it was found that UH-60s took significantly <u>less</u> time to do this than did MH-60Ks, and it was thought that this could have been an artifact of a greater number of UH-60s being disqualified due to mission failures. Consequently, a look at the main effect of helicopters on this variable would be enlightening, in view of the fact that for the present experiment, no one was disqualified for tardiness or for becoming lost.

Table 13 shows two very large main effects, one due to familiarity, ($\underline{M}$s= 852 sec, high; 1192, low) the other to helicopters ($\underline{M}$s=941, MH-60K; 1103, UH-60).

Table 13

Summary Source Table for Analysis of Variance for the Effects of Familiarity, Coordination, and Helicopters on Time to Acquire Checkpoint 9 (Acq[9])

Source	df	MS	F
Familiarity (F)	1	1,846,541	156.25a
Coordination (C)	1	11,637	<1.00
Helicopters (H)	1	417,801	35.35a
(F X C)	1	9,677	<1.00
(F X H)	1	513,193	43.43a
(C X H)	1	4,918	<1.00
(F X C X H)	1	4,778	<1.00
Error within Cell	56	11,818	

a. $\underline{p}$ <.001

Only the (Familiarity X Helicopters) interaction was significant, with MH-60Ks beating UH-60s when crews were unfamiliar ($\underline{M}$s= 1022 vs. 1362 sec), but with the MH-60K being edged out by the UH-60 (861 vs. 843) when crews were familiar. Both differences were significant via independent $\underline{t}$-tests (30 $\underline{df}$). For unfamiliar crews, $\underline{t}$= -6.40, $\underline{p}$ <.001; for familiar crews, $\underline{t}$= 2.41, p<.025.

These results in general seem to be supportive of Hypothesis 2, which called for greater differences in performance due to helicopter type under low than under high familiarity. It appears that the navigation system of the MH-60K makes more difference for unfamiliar than for familiar crews. This difference was probably muted in previous experiments because of the large number of UH-60 crews who failed to complete the mission. It is also interesting to note the significantly better performance on the terrain navigation tasks by familiar UH-60 crews. It is difficult to find a ready explanation for this difference. One very tentative hunch is that familiar MH-60K crews were able to miss checkpoints to a greater extent than were their UH-60 counterparts and still arrive at the final checkpoint, which may have added to the total time it took to acquire the final checkpoint. In short, time may have been of the essence for UH-60 crews to a greater extent than for those of the MH-60Ks, for whom the terrain flight navigation phase of the mission was much easier.

Although no helicopters were timed out of the mission, a dichotomous variable of (1=success, 0=failure) was used to distinguish those who took more than 1200 seconds to acquire the final checkpoint from those who took less than this amount of time. Means for this variable are expressed as percentages who succeeded. Table 14 provides another perspective on the interaction between helicopters and familiarity.

Table 14

Percentage of Mission Successes as a Function of Familiarity and Helicopter Type

	Helicopters	
	MH-60K	UH-60
Familiarity		
High	100	100
Low	100	19

It is obvious that, if the mission time out rule had been enforced for Experiment 4, approximately 60% of the UH-60s would have acquired Checkpoint 9 within the time allowed, as opposed to 100% of the MH-60Ks. All of the mission failures would have consisted of UH-60 crews who were low in mission and crew familiarity. Of these, approximately 80% would have had to abort their missions.

General Discussion

The present research effort must be deemed a success. Micro Saint, a modeling language which has been used extensively to develop predictive models of workload, has shown itself to be capable of modeling not only workload-driven crew errors, but also errors moderated by non-workload variables such as individual differences in crew communication strategies, experience, and mission familiarity. In brief, the model has demonstrated that crew-level error is not simply a workload problem, but one that is moderated by other individual difference variables.

Both helicopter variants of the model appear to have performed as expected, for the most part. The main expectations that were not confirmed by the model runs were the main effect of crew coordination strategy and the interaction between crew coordination and helicopter type for the terrain flight navigation phase of the mission. There were instead large main effects of helicopter type, with the MH-60K showing superior performance across the board for navigation, and of crew experience, with experienced crews generally outperforming their less experienced counterparts for this mission phase.

Although it was expected that highly coordinated crews in both helicopters would show similar performance in terrain flight navigation, it nevertheless appeared that well coordinated MH-60K crews outperformed well coordinated UH-60 crews. It appears that the present investigators were overly pessimistic about the enhanced navigation capabilities of the MH-60K. The workload-reducing cockpit of the newer helicopter makes it clearly superior in the area of NVG terrain navigation to its predecessor, even for crews whose coordination and communication is minimal. As a result, according to the model, it seemed that crews without previous mission experience and with low experience in the aircraft were much more likely to complete this very difficult and demanding mission in the MH-60K than in the UH-60.

Also, it appeared that when the variables of crew and mission familiarity and experience were manipulated independently, the latter variable affected terrain navigation performance for the MH-60K to a much greater extent than the former; in brief, it seemed that the technology of the MH-60K avionics minimized differences due to (crew and mission) familiarity. When familiarity and helicopter type were manipulated independently, as in Experiment 4, familiarity was seen to have a large effect on terrain flight navigation but not on approach and landing. Performance differences due to familiarity were also greater for the UH-60 than for the MH-60K, corroborating somewhat the findings of Experiment 3.

The generally superior performance of the MH-60K in NVG navigation is further illustrated by the large number of mission failures for the UH-60 versus the MH-60K, in Experiments 1, 2, and 3. Approximately 40% of the UH-60s failed to acquire the final checkpoint in time to initiate an approach and landing. Only 11% of the MH-60Ks in Experiment 2, and about 5% in Experiment 3, failed to complete a mission, and there were no other mission failures in the other experiments for this helicopter type.

The reader by now is probably wondering whether any data involving actual helicopters flying similar mission profiles exist. Although somewhat dated, the results of a research project by Wright and Gray (1964), provide a rare analysis of errors in actual terrain flight navigation practice missions. Wright and Gray investigated, through secondary analysis of previously-collected data, the likelihood that U.S. Army pilots and navigators flying daytime, nap-of-the-earth missions in H-13 helicopters would become geographically disoriented or have to abort their missions. These missions were either preplanned or immediately initiated (this could be roughly compared to high and low mission familiarity). The investigators concluded that geographical disorientation was exhibited on 52.5% of the preplanned and 64% of the immediately initiated sorties. This is roughly comparable to crews either straying too far off course or not being able to acquire critical checkpoints within a specified time frame; however, it is not possible to say whether disoriented crews would <u>always</u> have to abort their missions. Allowing for this and other caveats concerning comparability, it still appears that the estimates from the current modeling effort are not unrealistic at all. The investigators concluded that the Army NOE navigation techniques imposed excessive workload on the pilot, increasing the likelihood of geographical disorientation.

These techniques, using pilotage and paper maps, did not change radically between the time this research was done and the introduction of the UH-60 into service. The introduction of NVGs almost certainly exacerbated those problems identified by Wright and Gray. The results of the present simulation indicate that the hardware changes in the MH-60K may ameliorate many of the workload problems related to NOE navigation.

These results, taken together, may be important for the special operations mission scenario, for which the MH-60K was designed. Crews in this scenario would very unlikely to have flown over the same terrain before, and, if the current model has veridicality, it would seem to predict that even when this is the case, they would have a much better chance of succeeding at their mission than they would in the UH-60. While there are some areas that require further investigation (e.g., the greater sensitivity of obstacle clearance vs. navigation errors to different levels of crew coordination), the trends obtained from the parametric experiments were consistent with predicted performance differences, though those due to crew coordination did not attain conventional levels of significance for most of the critical navigation variables.

Although these initial examinations of the sensitivity of the model generally conform to expectations, there is much more that can be done with it. It should undergo more rigorous examination to determine whether it can account for more complex combinations of design- and crew-related variables.

Conclusions and Recommendations

The models for the two Blackhawk helicopter variants developed for this effort represent an exploratory attempt to predict certain types of crew error as a function of key variables, including crew coordination strategies, workload levels, experience, familiarity, and equipment design. While they take a first step toward predicting error in a complex system such as a helicopter, there is more that can be done in future efforts.

They could serve as the foundation for a family of models used to predict different kinds of crew error in helicopters. These models specifically predict errors related to NVG navigation over terrain and later through trees during a low level approach to a confined landing zone. Follow-on models could be developed which predict other types of error such as obstacle clearance errors during all phases of NOE flight, errors in flight control manipulation, or decision making errors. It would also be logical to expand the applicability of the models to other operational missions and flight regimes such as attack scenarios, sustained operations scenarios, multi-aircraft operations, or alternate equipment configurations. In short, this effort represents a first step in the direction of using simulation modeling tools to assess the safety and mission effectiveness of Army aviation.

Finally, the reader should be made aware of the major limitation of the present research effort. The input data from the model have not been empirically validated against an external criterion. In short, the level of validation thus far has been internal, with the goal of determining if predictions from the model are internally consistent. Thus the reader is cautioned not to draw too detailed conclusions from the preceding set of analyses with the model. This would suggest a need to validate the model externally against a similar mission profile in a full-mission simulator, with analogous performance measures, and crews varying in coordination strategy, familiarity, and experience.

REFERENCES

Alkov, R. A. (in press). Attitude change in Navy/Marine flight instructors following an aircrew coordination training course. International Journal of Aviation Psychology.

Bierbaum, C.R., Szabo, S.M. & Aldrich, T. B. (1989). Task analysis of the UH-60 mission and decision rules for developing a UH-60 workload prediction model. ARI Research Product 89-08 AD A210763.

Bierbaum, C.R., and Hamilton, D.B. (1990). Task analysis and workload prediction model for the MH-60K mission and a comparison with UH-60A workload predictions. ARI Research Report 1576.

Boeing Commercial Airplane Company (1985). Statistical summary of commercial jet aircraft accidents, worldwide operations, 1959-1984.

Chidester, T.R. and Foushee, H.C. (1988). Leader personality and crew effectiveness: Factors influencing performance in full-mission air transport simulation. Proceedings of the 66th Symposium of the Aerospace Medical Panel on Human Behavior in High Stress Situations in Aerospace Operations, Advisory Group for Aerospace Research and Development, The Hague, Netherlands, pp. 7-1-7-9.

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

Helmreich, R.L., Chidester, T.R., & Foushee, H.C. (1989). How effective is cockpit resource management training? Issues in evaluating the impact of programs to enhance crew coordination. NASA Draft Technical Report 89-2.

International Civil Aviation Organization (1984). Accident prevention manual (1st. Ed.) ICAO Doc 9422-AN/923.

Laughery, K.R., Drews, C., Archer, R., & Kramme, K. (1986). A Micro Saint simulation analyzing operator workload in a future attack helicopter. Proceedings of the National Aerospace Electronics Conference (NAECON), Vol. III, pp. 896-903.

Rasmussen, J. (1981). Human errors: A taxonomy for describing human malfunctions in industrial installations. RISO National Laboratory, Denmark, 1981.

Reason, J.T. (1981). GEMS: Framework for locating common human error forms. Human Performance Models for Computer-Aided Engineering. NASA.

Stewart, J.E. and Lofaro, R.J. (1990). A meta-analytic approach for relating subjective workload assessments with U.S. Army Aircrew Training Manual (ATM) ratings of pilot performance. ARI Technical Report 913, AD A230 127.

Wright, R. H. and Debonis, J.N. (1975). Object visibility patterns in low level flight. U.S. Army Aeromedical Research Laboratory, USAARL Report No. 76-5. AD A016886

Wright, R.H. and Gray, T.H. (1964). Army low altitude navigation: System considerations and procedural solutions. Proceedings of the Tenth Army Human Factors Research and Development Conference, 246-253.