AL/HR-TP-1997-0001

UNITED STATES AIR FORCE ARMSTRONG LABORATORY

A COMBAT MISSION TEAM PERFORMANCE MODEL: DEVELOPMENT AND INITIAL APPLICATION

Denise R. Silverman

V. Alan Spiker

Steven J. Tourville

HUGHES TRAINING, INC.
Training Operations
6001 South Power Road, Building 560
Mesa AZ 85206-0904

Robert T. Nullmeyer

HUMAN RESOURCES DIRECTORATE
AIRCREW TRAINING RESEARCH DIVISION
6001 South Power Road, Building 558
Mesa AZ 85206-0904

June 1997

AIR FORCE MATERIEL COMMAND
ARMSTRONG LABORATORY
HUMAN RESOURCES DIRECTORATE
7909 Lindbergh Drive
Brooks AFBTX 78235-5352

NOTICES

This report has been reviewed and is approved for publication.

ROBERT R. NULLMEYER
Project Scientist

DEE H. ANDREWS
Technical Director

LYNN A. CARROLL, Colonel, USAF
Chief, Aircrew Training Research Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
June 1997

3. REPORT TYPE AND DATES COVERED
Final - September 1995 to December 1996

4. TITLE AND SUBTITLE
A Combat Mission Team Performance Model: Development and Initial Application

5. FUNDING NUMBERS
C - F41624-95-C-5011
PE - 62202F
PR - 1123
TA - B2, B3
WU - 06, 01

6. AUTHOR(S)
Denise R. Silverman
Steven J. Tourville
V. Alan Spiker
Robert T. Nullmeyer

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Hughes Training, Inc.
Training Operations
6001 South Power Road, Building 561
Mesa, AZ 85206-0904

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Armstrong Laboratory
Human Resources Directorate
Aircrew Training Research Division
6001 South Power Road, Building 558
Mesa, AZ 85206-0904

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AL/HR-TP-1997-0001

11. SUPPLEMENTARY NOTES
Armstrong Laboratory Technical Monitor: Dr Robert T. Nullmeyer, (602) 988-6561. This technical paper has been published in the Proceedings of the 18th Interservice/Industry Training Systems and Education Conference, held 2-4 December 1996 in Crlando FL.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
Aircrew coordination is now an integral part of all Air Force combat mission training. A vast body of literature exists that deals with aircrew coordination, a subset of which addresses combat mission training. While it is commonly assumed that effective aircrew coordination leads to improved mission performance, surprisingly few studies have demonstrated an empirical link between them using tactically realistic combat scenarios. We present a conceptual model of team performance measurement in which aircrew coordination, team performance, mission performance and their interrelationships are operationally defined. The model builds on the seminal study conducted by the Air Force in 1989 and provides a useful framework for interpreting crew resource management research from other laboratories. Validation of the model has begun with Air Force Special Operations Command (AFSOC) MC-130P aircrews and preliminary data are provided that support key elements of the model.

14. SUBJECT TERMS
Aircrew; Aircrew coordination; Combat mission; Combat mission Training; Coordination; Crew resource management; Flight training; Team performance; Training;

15. NUMBER OF PAGES
16

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION ABSTRACT
Unclassified	Unclassified	Unclassified	UL

PREFACE

This report documents a paper which was presented at the Interservice/Industry Training Systems and Education Conference (I/ITSEC) which was held in Orlando FL from 2-4 December 1996. This research was conducted under Work Unit 1123-B2-06, Aircrew Training Research Support; Contract F41624-95-C-5011 with Hughes Training, Inc-Training Operations; and 1123-B3-01, Special Operations Forces (SOF) Aircrew Training and Mission Preparation Research. The Laboratory Principal Investigator was Dr Robert T. Nullmeyer; the Laboratory Contract Monitor was Mr Daniel Mudd.

This effort is part of an Armstrong Laboratory, Human Resources Directorate, Aircrew Training Research Division (AL/HRA) program to help identify how mission rehearsal devices can be most effectively used in mission rehearsal, and to assess the actual impact of this technology as it becomes available in the Air Force SOF community.

A COMBAT MISSION TEAM PERFORMANCE MODEL: DEVELOPMENT AND INITIAL APPLICATION

Denise R. Silverman
Hughes Training, Inc. - Training Operations
Albuquerque, New Mexico

V. Alan Spiker
Anacapa Sciences, Inc.
Santa Barbara, California

Steven J. Tourville
Hughes Training, Inc. - Training Operations
Albuquerque, New Mexico

Robert T. Nullmeyer
Armstrong Laboratory
Aircrew Training Research Division
Mesa, Arizona

INTRODUCTION

During the past 20 years, cockpit resource management (CRM) has become a widely used component of aircrew training programs for both the civil and military communities (Gregorich & Wilhelm, 1993). In a landmark study of the effects of workload on aircrew performance, Ruffell Smith (1979) reported that the behaviors that most differentiated effective crews from weaker ones were in the areas of leadership, decision making, and resource management, setting the foundations for training "softer skills." It was believed by researchers, although not fully accepted by crewmembers, that training in these areas would yield large dividends in terms of increased flight safety, more evenly distributed crew workload, and more efficient communication.

Despite CRM's acknowledged importance, there is little evidence directly linking effective crew coordination and good mission performance, particularly in the context of combat mission training (CMT). We begin our paper by discussing three critical CRM issues: (a) content of CRM training, (b) measuring CRM effectiveness, and (c) collecting CRM data. Next, we describe the key elements of a comprehensive, measurement-oriented conceptual model of team performance. It is our contention that the lack of such a model has stifled previous research efforts in this area. We then use the model to interpret several CRM studies. We conclude by describing some initial results from applying the model to MC-130P Annual Refresher Training being conducted by the Air Force Special Operations Command (AFSOC) at the 58th Training Support Squadron, Kirtland AFB, NM.

Content of CRM Training

Early CRM training programs lacked definition and delineation of the required aircrew coordination behaviors in operational terms. The behaviors trained were often too "touchy-feely" to either develop concrete measures of these behaviors or to be accepted by aircrews as necessary for flying the aircraft. Typical CRM topical areas included communication processes, team building, and workload management (Gregorich & Wilhelm, 1993). It is not immediately clear what the reinforceable or observable behaviors associated with these areas would be. It is also possible that this training might not be effective in a more tactically oriented, CMT setting, as its focus on non-technical areas would be overshadowed by tactical and combat skill requirements.

In addition, early CRM training programs lacked a larger team training approach. In both the commercial airline industry and the military, the initial focus was on individuals rather than the entire aircrew team. The first CRM course in the early 1980s attempted to improve the attitudes of individual pilots in order to promote more communication and information-sharing in the cockpit. CRM training was designed to impact selected aircraft commanders (ACs) to "fix" those most likely to resist information from co-pilots or other crewmembers in time-critical, high workload situations (Helmreich, 1995). Similarly within the USAF, CRM training was originally focused on individual ACs. Over time, coordination concepts have been expanded to include other crewmembers, and ultimately some envision it encompassing the entire combat mission team, including intelligence, weapons and tactics, logistics, weather, airborne command and control, air traffic control, maintenance, and the ground "customers" supported by the aircrew (Andrews, Bell, & Nullmeyer, 1995).

CRM training programs were also initially quite generic, where all airlines and airframe types received similar training. More recently, some airlines have taken a more problem-oriented approach in which each airline explores the particular CRM-related problems that plague their operations rather than taking a global, industry-wide perspective. Likewise, the USAF has moved toward weapon system-specific simulator training and the corresponding administration and assessment of CRM course materials.

The precise delineation of CRM principles that can be tied to operationally relevant behaviors, the appropriate team training approach, and context-specificity are important CRM training program components. Including these factors in CRM training programs will: (a) enable researchers and training specialists to provide feedback and reinforcement on specific CRM behaviors to particular crewmembers, (b) increase crewmember and instructor motivation toward learning and applying CRM principles, and ultimately, (c) establish an environment for determining CRM training effectiveness.

Measuring CRM Effectiveness

It is encouraging that some researchers have been successful in linking CRM with mission performance. However, this has mostly been accomplished using fairly narrow measures of CRM, such as communication or attitudes.

In an early study of crew coordination, Krumm and Farina (1962) investigated the impact of integrated simulator training on B-52 mission effectiveness. They collected process data on the pattern and rate of communication between crewmembers during selected segments of the training mission. They also collected objective measures of performance, including navigational and bombing accuracy. They found that the method of training had a positive impact on coordination, as the crews who trained together had better patterns of communication than the crews who did not. Second, and most important, they noted that the quality of the communication/coordination patterns was significantly related to both navigation and bombing accuracy. For example, crews who navigated more accurately also volunteered more information.

More recently, Predmore (1991) examined communication patterns associated with a number of major airline accidents. Overall, he found that the rate of thought units expressed in the cockpit increased dramatically following a stressful event. In one accident, the crew expressed an average of five thought units per minute prior to loss of the cargo door and almost 19 units per minute following door loss. Further analysis revealed that in unsuccessful responses to emergencies, cockpit communication patterns develop in which one crewmember dominates, insufficient communications are transmitted/received from outside the cockpit, and a large percentage of communications are either interrupted or incomplete.

Schmidt (1987), in her work on C-130 aircrews, also found several communication patterns that were associated with more successful crews. These included having a greater frequency of communication, fewer communications "left open," and use of problem solving as a primary source of conflict resolution. Both of these examples give some indication that effective crew coordination is related to mission effectiveness. However, we feel that they have only captured a small subset (i.e., communication) of the relevant crew coordination processes that affect crew performance, especially in terms of CMT.

Besides communication, other researchers have attempted to establish links between CRM training and subsequent attitude change. For example, after implementing a revamped Aircrew Coordination Training (ACT) program, the Army initially assessed attitude change. Surveys showed that aviators and instructor pilots (IPs) exhibited positive changes in attitudes toward ACT over the course of training (Zeller, 1992). While encouraging for implementation, the results did not pinpoint the locus of ACT program impact. That is, it was not clear whether the program was affecting crew performance or mission outcome. In a similar vein, airline researchers (Gregorich & Wilhelm, 1993) have noted that whereas CRM seminars have demonstrated an "increase in targeted attitudes and motivations toward

CRM concepts, there have been no links to behaviors in LOFT [line-oriented flight training] or to flight operations—i.e., mission performance" (p. 193).

Although these response measures say quite a bit about CRM, they do not "tell the whole story," particularly as it pertains to CMT. As discussed later in our conceptual model, a key aspect of CRM involves the measurement of tactical behaviors and processes, and their likely links to mission performance. In addition, for determining definitive relationships between CRM processes and performance, one should not rely solely on any single measure. Our approach, therefore, measures CRM along multiple dimensions that encompass reinforceable behaviors and individual attitudes.

Collecting CRM Data

CRM data can be collected using a variety of methods, such as audiovisual recordings, questionnaires, and direct observations. While a rich source of data, analysis of audiovisual recordings is highly labor-intensive. A primary advantage of using videotape is the ability to review mission sessions repeatedly (e.g., Schmidt, 1987), however, this can be offset by logistical problems. In our setting, we have observed such problems as low light levels under NVG conditions that make the images blurry and the lack of digital video display times to facilitate analysis of mission events during tape review (Silverman, 1995). From an operational perspective, relying solely on videotape to collect crew coordination data is highly impractical if one goal of the data collection effort is to transfer the techniques for use in the actual aircraft.

Questionnaires, too, are useful, but any prespecified set of questions can miss potentially relevant crew coordination information. That is, in order to ensure their completion and make them reasonable lengths, they will almost certainly have a limited scope. Similarly, untrained observers can miss critical information because they lack the insight gained from experience and training that is required to recognize exceptional or weak crew performance, and the behaviors that could be used to discern relationships between coordination and performance.

Independent assessments of team coordination processes and performance are essential for avoiding artificially inflated correlations that are inevitable when obtaining these measures from the same rater. As an example, the airlines recently recognized that the next step in their CRM research entails demonstrating positive correlations between CRM processes and crew effectiveness. Law and Wilhelm (1995) found that certain crew coordination behaviors were, indeed, related to overall crew effectiveness. Unfortunately this may have been unduly confounded because they used the same evaluators to provide team coordination process and performance judgments.

The most effective approach for data collection will use a combination of these methods. Thus, trained observers will capture salient CRM behaviors during CMT using customized checklists, rating sheets, and surveys. When available, videotapes will be consulted to substantiate observations and support debriefings.

A CONCEPTUAL MODEL

We believe that the conduct of CRM research within a highly turbulent CMT environment should be guided by a comprehensive measurement framework. Such a model would provide a common language to define the content of CRM training, establish valid indices to gauge CRM impact, and specify appropriate procedures for collecting CRM data. In the following section, we present a conceptual model (see Figure 1) currently used to guide our research (Spiker, Tourville, Silverman, & Nullmeyer, 1995).

In Figure 1, the concepts and the arrows linking them flow from left to right, reflecting an implicit timeline (arrival through outbrief) of CMT activities. Three modules feed into Team Coordination Processes. The first two, Crew Background and Baseline Attitudes, reflect the fact that, going into training, aircrews will vary in terms of their background experiences (e.g., squadron affiliation, hours flown together as a crew) and attitudes toward CRM principles and training. With regard to the latter, analysts have posited that crew attitudes towards CRM may set an upper limit on the amount of positive change that can be expected to result from CRM training (Wilhelm, Roithmayr, & Helmreich, 1992). To capture these factors, we administer a crew background survey (the oval labeled CBS) and a pre-Team-Mission Attitudes Questionnaire (the oval labeled TMAQ1).

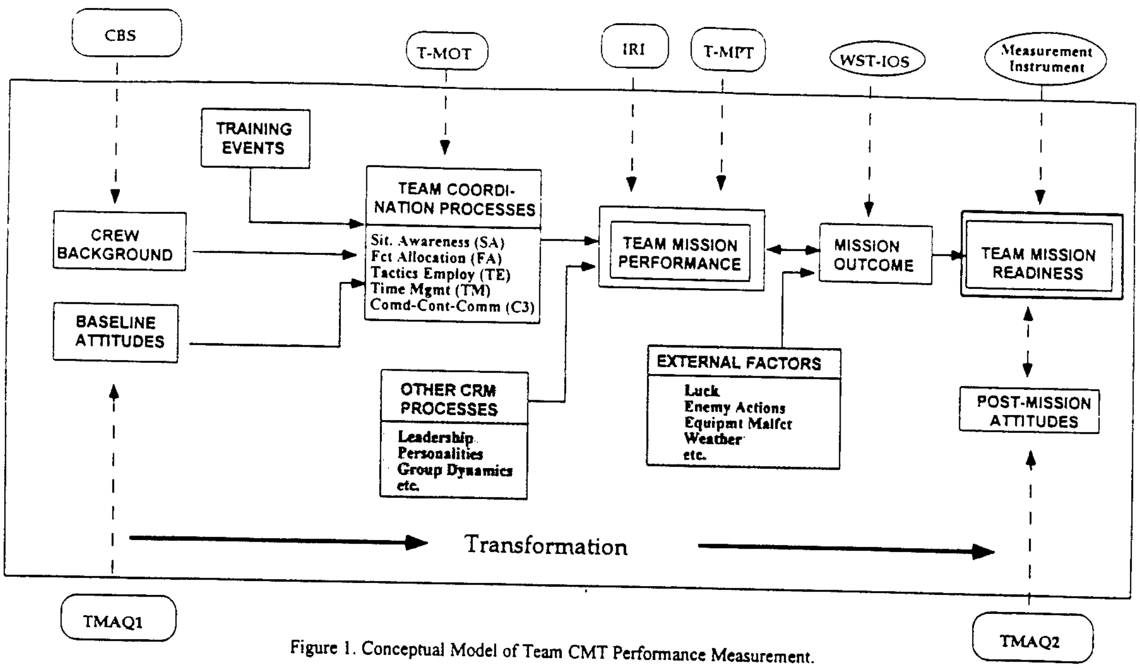

Figure 1. Conceptual Model of Team CMT Performance Measurement.

We represent Training Events as a single, undifferentiated component. However, we recognize that, in practice, there are many salient training events during CMT. These include the CRM academics the crew receives, all relevant technical and combat tactics training, as well as the CMT mission scenario and scripted events that are presented before and after the crew flies in a simulator or an aircraft.

Within the Team Coordination Process module, we include five team coordination processes. These processes were selected for further exploration in our research based on their: judged relevance to the AFSOC mission environment, appropriateness to the high levels of experience and motivation of many MC-130P aircrews, applicability to CMT, and amenability to measurement by outside observers. Where possible, we attempted to identify functional areas that make contact with the CRM dimensions that have been identified by other researchers.

The five team coordination processes are: (a) Function Allocation (FA)—the division of crew responsibilities so that workload is distributed among the crew, avoiding redundant tasking, task overload, and crewmember disinterest or noninvolvement, and where tasks are allocated in such a manner that crewmembers are able to share information and coordinate responsibilities; (b) Tactics Employment (TE)—all analytic activities necessary to avoid or minimize threat detection or exposure, and to successfully coordinate complex mission events and multiple mission objectives; (c) Situation Awareness (SA)—maintenance of an accurate mental picture of mission events and objectives as they unfold over time and space; (d) Command-Control-Communications (C3)—those activities required to involve external parties in the mission and to maintain communications with these external team members, communication within the crew, and controlling the sequence of mission events according to the mission execution plan; and (e) Time Management (TM)—the ability of the combat mission team to employ and manage limited time resources so that all tasks receive sufficient time to be performed correctly, and critical tasks are not omitted.

These five areas are depicted in white in Figure 1 to indicate their coverage by our Team-Mission Observation Tool or T-MOT (Tourville, Spiker, Silverman, & Nullmeyer, 1996) represented by the oval above this module. Since we know that these five areas do not encompass the entire domain of what would properly be considered team coordination, we represent Other CRM Processes in the gray-shaded box, feeding into Team Mission Performance. These other processes may influence team mission performance, but they are currently not measured in our research.

The output of the Team Coordination Processes module feeds directly into the Team Mission Performance module. By Team Mission Performance, we mean those indices that directly result from the successful (or failed) execution of important team processes (e.g., TE, FA). In our research, Team Mission Performance is reflected in such indices as quality of the pre-mission briefing, completeness of the navigation chart(s), as well as instructor-supplied ratings of how well the team as a whole executed each phase of the mission, including mission preparation, low level navigation and threat avoidance, aerial refueling (AR), air drop, and infil/ exfil. The ovals above Team Mission Performance refer to the two tools (Instructor Rating Instrument or IRI and the Team-Mission Performance Tool or T-MPT) that we are using to collect this information.

As conceptualized here, Mission Outcome consists of those indices that would be used to conclude that the team's mission accomplished its stated objectives. When the mission is performed in a weapon system trainer (WST), these can often be recorded by computer. Example outcomes include accuracy of airdrops, performing infils within prescribed control time windows, time spent on ground before an exfil, and minimizing (or avoiding) exposure to threats.

Mission Outcome has the advantage of encompassing the criterion environment and being the ultimate yardstick in which the operations and training communities are most interested. But from a team research standpoint, reliance on outcome as the sole effectiveness criterion is risky as there are many external factors that may act to degrade outcome, but have little or nothing to do with the effectiveness of the combat mission team. Some of these External Factors are listed in the gray-shaded box underneath Mission Outcome (e.g., luck). Unfortunately, when assessing team coordination under operational or training conditions, researchers usually have little control over these external factors. As such; researchers inherit a great deal of noise and uncontrolled variability in their outcome measures, making it difficult to infer whether good or poor team coordination has occurred.

We connect Team Mission Performance and Mission Outcome with a bidirectional arrow to indicate that some of our indices of team performance are based on data that would normally be used to measure Mission Outcome. That is, we contend that performance leads to interim outcomes which in turn impact subsequent team mission performance. For example, crew coordination affects the quality of the mission execution plan that is produced which in turn affects subsequent mission performance.

The Team Mission Readiness module reflects the end state of the crew at the conclusion of CMT. Following the execution of the mission scenario in the WST and formal debriefing by the instructors, the aircrew should have become "transformed" into a team which is ready to perform similar operational missions. While such transformations are not observed directly, they can be inferred from behavior changes noted by trained instructors as well as by tracking how well the crew performs once it returns to its operational unit. We have placed a Measurement Instrument oval above this module to reflect our belief that one can measure an aircrew's mission readiness. In our current research, this is not being measured; however, we plan to pursue this in the future.

The last module is Post-Mission Attitudes. By comparing crewmembers' attitudes toward CRM after academic training and the tactically rich mission-oriented simulator training (MOST) mission (the post-Team Mission Attitudes Questionnaire or TMAQ2 oval) with those obtained during the pre-training baseline, researchers will be able to determine the degree to which crewmembers' attitudes changed over the course of CMT. From a strict experimental standpoint, one cannot unequivocally attribute a change in attitude to the occurrence of CRM-oriented training. Nevertheless, a close association in time and space between CRM training and attitude change is certainly suggestive of a direct link, and is an assumption that is shared by notable CRM experts (Helmreich, Merritt, Sherman, Gregorich, & Weiner, 1993). It is important that one not mistake a relationship between Training Events and CRM Attitude Change for the more fundamental relationship between Team Coordination Processes and Team Mission Performance. It is this latter relationship that has the greatest implications for CMT.

We feel the model provides several key elements that are unique and critical for establishing crew coordination process and performance links. They are: (a) multiple measures of crew performance, beyond mission outcome; (b) assessment of multiple coordination processes; (c) measures of crew coordination process and performance throughout the timeline of training; (d) baseline measures of individual acceptance of crew coordination training and principles; and (e) independent assessments of crew coordination process and performance.

USING THE MODEL TO INTERPRET CRM MILITARY RESEARCH

One useful application of our model is to organize and interpret past CRM research within the military. A few examples are discussed below.

Army

A study by Thornton, Kaempf, Zeller, and McAnulty (1992) examined the relationship between aircrew coordination and mission effectiveness. Nineteen crews of two aviators each performed a combat-oriented mission in an advanced UH-60 Black Hzawk simulator. Several hours of planning time were given to each crew prior to the mission. Video recordings were taken during mission execution.

As viewed through our model, Thornton et al. took the following approach. Team coordination was measured exclusively in terms of communication. To that end, two researchers worked independently to develop a communications protocol that assessed each crew in terms of the rate, pattern, content, and quality of interactions along 13 functional categories (inquiry, command, declarative, etc.). For team performance, mission effectiveness was defined in terms of three general functions—navigation accuracy, threat evasion, and shooting a nonprecision instrument approach. These were chosen based on their strong a priori relationship to mission success. Navigation performance was measured in terms of course deviations from the planned ground track and amount of time spent off-course. Threat avoidance performance was measured as the number of threats encountered during the mission and time exposed to each threat. The quality of the crew's instrument approach performance was rated by two researchers using a detailed checklist derived from the supplied instrument approach plate.

Overall, there was some evidence that crew coordination processes, as defined by patterns and types of cockpit communications, showed a significant relationship to some of the mission effectiveness indices. For example, the researchers reported that crews who were successful in evading threats had a pilot-flying (PF) who issued more acknowledgments than his PF counterpart in the unsuccessful crews. Although rate of communication did not differentiate among the crews who performed poorly (i.e., those who navigated inaccurately, were exposed to threats, and had poor approach proficiency) and those who did well, there were trends in the data which suggested that certain types of communications profiles were consistently related to outcome.

In summarizing their results, the authors conclude that there is some evidence for a direct relationship between the communications aspect of aircrew coordination and outcome-based measures of performance. Importantly, they noted low levels of technical proficiency observed in many of the crews. Indeed, problems with map interpretation, terrain feature identification, and issuing standard radio calls were prevalent among the less successful crews. Their results must be interpreted in this light because these skills should be mastered prior to learning CRM skills.

In addition, the researchers' definition and measurement of coordination was, from our standpoint, a bit narrow since many other relevant subprocesses (SA, resource allocation) were not included. Interestingly, the authors provided anecdotal evidence to support a relationship between planning quality and mission outcome. They noted that crews who performed the instrument approach better had spent more time studying the approach plates during planning, and hence needed to refer to it less often during the high workload landing phase. Such evidence encourages our view that measurements of team coordination processes should encompass the entire mission, from mission planning through mission completion.

Joint Forces

A study by Dwyer, Fowlkes, Oser, and Salas (1996) is consistent in many ways with our conceptual model of team performance. The expressed purpose of the study was to develop a performance assessment technique for distributed interactive simulation (DIS) environments. In laying out their technique, Dwyer et al. clearly delineate between process and outcome measures of team performance, a distinction our model of team performance requires. They support this delineation by using two tools, the Targeted Acceptable Responses to Generated Events or Tasks (TARGETS) and the Teamwork Observation Measure (TOM), to capture team coordination processes, and one tool, the Unit Performance Assessment System (UPAS), to capture outcome measures of team performance (e.g., number of enemy vehicles damaged). They recognize that "the development and use of multiple performance assessment tools provide an opportunity to examine performance from different perspectives [which] paint[s] a detailed picture of performance" (p. 372)--perhaps the cornerstone of our model.

Dwyer et al. examined the usability and reliability of TARGETS and TOM across five nodes of the Multi-service Distributed Training Testbed (MDT2) simulation. They had a team of observer/controllers (OCs) observe DIS training for Close Air Support (CAS) missions over five days. The OCs were responsible for completing the TOM and TARGETS and to role-play higher echelon positions as necessary during the course of the DIS training. Across the three phases of the mission they found that: (a) the instruments were reliable, as OCs at different training nodes (i.e., from different services) had similar response patterns on both TOM and TARGETS; (b) TOM and TARGETS showed team coordination process learning trends from day 1 to day 5; and (c) each tool provided valuable performance feedback information for the trainees in after action reviews.

Regarding the learning trends over days 1-5, Dwyer et al. found differences between the learning curves plotted for TARGETS and those for TOM. Although both instruments revealed performance improvements over days 1-5, the rise was more subtle for TOM assessments of the contact point and attack phase. We point this out primarily because of the inference they draw from this finding. That is, they describe both TARGETS and TOM as tools to measure the team coordination process, but in fact, their data reveal that the two instruments may be measuring two different aspects of team process. The TARGETS instrument may assess the presence or absence of a process, whereas TOM may provide some qualitative account of the process. Apparently, "the tasks that should have been performed were performed, however how well they were performed with respect to the teamwork dimensions could have been enhanced" (p. 377).

As seen through our conceptual model, this study provides a rather complete picture of team performance within a combat mission environment. One of the truly admirable features of this study is its examination of team coordination in a much larger combat team environment. Other praiseworthy aspects of this study include: on-line data collection with highly trained observers, collecting team process and performance data from separate sources, collecting data through mission phases, and using a realistic tactical scenario. However, as seen through our conceptual model, Dwyer et al. fall short in one area. That is, they do not report any performance data to allow exploration of the fundamental team process/performance relationship.

Air Force

Povenmire, Rockway, Bunecke, & Patton's (1989) study of B-52 aircrew coordination represents one of the strongest attempts to demonstrate a direct relationship between crew coordination processes and mission performance. This study employs a number of the methodological features we have discussed in the presentation of our model. These features are particularly valuable for examining process-outcome correlations in the context of CMT.

Povenmire et al. observed seven intact aircrews fly a complex, tactically realistic mission scenario in a high fidelity B-52 WST. The scenario entailed conducting a long-range bombing mission requiring the penetration of enemy threats, accurate dropping of bombs, and intricate navigation and maneuvers. Highly trained CRM evaluators assessed aircrew coordination and mission performance with separate sets of raters used for each measure. Mission performance was evaluated based on three factors: bombing accuracy, threat avoidance, and technical skill. The latter factor consisted of a number of subfactors, such as maintaining appropriate altitude, performing accurate navigation, and staying within designated control times. The researchers asked the evaluators to rank order the crews from best to worst, based on their subjective impressions of the three mission performance factors.

The primary analysis assessed the correlation between overall aircrew coordination and the crew's mission performance ranking. A strong positive relationship ($r = .84$) was obtained. Povenmire et al. then compared the experts' ratings of mission performance with the individual mission outcome factors. Part-whole correlations showed that the experts primarily used bombing accuracy to make their overall judgment of mission performance, as evidenced by a significant correlation of $r = .81$. The researchers also performed a series of part-whole correlations on the coordination data to determine the skill dimensions that had the highest loadings. These included practicing inquiry and advocacy, avoiding distractions, distributing workload, and resolving conflicts.

Despite the simplicity of its design and data analysis strategy, the Povenmire et al. study stands as one of the most clear-cut demonstrations of the relationship between aircrew coordination processes and mission performance. Indeed, the elegance of their design is somewhat deceptive in terms of providing unusually clear insights regarding the coordination subprocesses that best predict mission performance. We capitalize on the Povenmire et al. approach and expand upon it in our initial application of our conceptual model of team performance measurement.

CURRENT RESEARCH

We have recently begun to validate our model with MC-130P crews (an AC, a copilot, two navigators, a flight engineer, and a communication systems officer) attending week-long Annual Refresher Training at Kirtland AFB, NM. The five-day curriculum provides both CRM academic training at the beginning of the week and a tactically rich MOST mission at the end of the week with multiple technical training events for each crew position in between. Briefly, and in accord with our model (Figure 1), we are collecting pre- and post-CRM training attitude assessments (TMAQ1 and TMAQ2) and crewmember background information. During the MOST mission we collect: (1) instructor ratings (IRIs) of crew and individual performance across phases of the flight, including mission planning, low level, AR, airdrop and infil/exfil; (2) coordination assessments of the crewmembers and crew across the five identified subprocesses and phases of flight (T-MOT); (3) crew mission performance assessments across phases of flight (T-MPT); (4) as well as outcome measures tallied from observations and WST-IOS printouts. Independent raters make the assessments of crew coordination and mission performance. The primary goal of our research is to determine team process and performance links. We feel that the application of our model will enable us to do this by overcoming many of the shortfalls of past research and allowing us to more completely characterize good versus poor crews.

Although data collection is far from complete, early indications are that some of our assertions about the utility of the model are correct. For example, in some of the past CRM research, there have been failures to demonstrate team process and performance relationships because researchers equated mission outcome and mission performance, thereby restricting performance variation between crews. In structured observations of just four crews, we are already noting discriminations between strong and weak crews that could not have been made based on mission outcome alone. That is, we have seen one outstanding crew, two average crews, and one crew who had difficulty, yet all successfully completed the MOST mission. We have been able to draw out distinctions between the crews primarily because we include detailed mission performance measures as a part of our strategy.

Initial results also indicate that having independent observers make process and performance assessments across phases of flight is proving quite fruitful allowing us to isolate process/performance relationships. For example, early indications suggest that team coordination planning process ratings are highly correlated ($r = .87$) with performance ratings of mission planning (average product scores). So far, overall assessments of team coordination are most related to quality of mission planning products: the best crew observed had the highest overall planning performance score and the highest mission planning process rating. Other characteristics of this crew, from a process level, are quite interesting and hopefully will be supported with observations of additional crews. The crew's high overall assessment of coordination seemed to be influenced primarily by the crew's superior SA and TE, as crewmembers across phases of flight were rated as 4s and 5s in these areas. This is in contrast to the weaker crew observed which had several crewmembers, most notably the AC and the Right Navigator, rated as 2s and 3s on SA and TE across several mission phases.

IMPLICATIONS

These are very preliminary findings, but they are suggestive of crew process/performance relationships that we hope to unveil as we observe more crews. We fully expect that ideas for a number of procedural improvements to CMT should accrue from our research which can be folded back into the Annual Refresher Training course that is serving as our research testbed. Areas where procedural enhancements might be found include:

Instructor reinforcement of key behaviors. We will likely identify ways that instructors can provide effective feedback to crews, particularly in terms of providing immediate reinforcement for positive CRM behaviors across our five subprocesses (FA, TE, SA, C3, TM). For example, we have seen that when crews explicitly designate specific crewmember mission duties (FA) early in the planning session, they perform better during mission preparation. This may be a behavior that the instructor can prompt or praise during planning sessions.

Selective cross training. Specialized cross training in key crew functions (e.g., threat identification, navigation updating) might be observed as a way to improve FA during high workload periods of the mission. This may be particularly important for the FE and CSO. We have already observed several crews that underutilized these two crewmembers. It may be possible to provide them with additional, more interactive and nontraditional roles during planning that will increase the crew's resource efficiency and preparedness for their mission.

Team training. Our results may also indicate that having the crew spend more time together during CMT is important and we will specify the nature of resulting improvements. Presently, much of CMT is done on a duty position-specific basis, leaving crews little time to solidify the CRM concepts they learn early in the week of training. We may be abl determine ways to increase ensemble crew trai by focusing on areas where the benefits of CMT h been demonstrated.

Instructor acceptance of CRM. Techniques that better engage CMT instructors in the CRM process and that help instructors accept and embrace CRM ma be an additional outcome. Following the week's Annual Refresher Training and our data collection, we have been provided with five to ten minutes of "debrief" time where we are given the opportunity to explain our research efforts. This has been very well received, especially by the instructors, who seem to appreciate the strong connections we provide between our five CRM subprocesses and the ta al environment. We may be able to package this f after the research is completed.

REFERENCES

Andrews, D. H., Bell, H. H., & Nullmeyer, R. T. (1995). team training: A tutorial. From *the Seventeenth Inter el Industry Training Systems and Education Conference.* Albuquerque, NM.

Dwyer, D. J., Fowlkes, J., Oser, R. L., & Salas, E. (1996). Case study results using distributed interactive simulation for close air support training. In *Proceedings of the International Training Equipment Conference.*

Gregorich, S. E., & Wilhelm, J. A. (1993). Crew resource management training assessment. In E. Wiener, B. Kanki, & R. Helmreich (Eds.), *Cockpit resource management* (pp. 173-198). San Diego, CA: Academic Press, Inc.

Helmreich, R. L. (1995). Interviews conducted at University of Texas, Austin, September 26, 1995.

Helmreich, R. L., Merritt, A. C., Sherman, P. K., Gregorich, S. E., & Weiner, E. L. (1993). *Flight management attitudes questionnaire (FMAQ).* (NASA/UT/FAA Technical Report 93-4). Austin, TX: University of Texas.

Krumm, R. L. & Farina, A. J. (1962). *The effectiveness of a B-52 integrated flight simulator for its coordination training potential as measured by crew communications and performance measures* (AIR-327-61-FR-239). Washington, DC: American Institute for Research.

Law, J. R., & Wilhelm, J. A. (1995). Ratings of CRM skill markers in domestic and international operations: A first look. In *Proceedings of the Eighth International Symposium on Aviation Psychology.* Columbus, OH.

Povenmire, H. K., Rockway, M. R., Bunecke, J. L., & Patton, M. W. (1989). *Evaluation of measurement techniques for aircrew coordination and resource management skills* (UDR-TR-89-108). Williams AFB, AZ: Air Force Human Resources Laboratory, Operations Training Division.

Predmore, S. C. (1991). Microcoding of communications in accident investigation: Crew coordination in United 811 and United 232. In *Proceedings of the Sixth International Symposium on Aviation Psychology.* Columbus, OH.

Ruffell Smith, H. P. (1979). *A simulator study of the interaction of pilot workload with errors, vigilance, and decisions* (NASA TM-78482). Moffett Field, CA: NASA-Ames Research Center.

Schmidt, K. (1987, draft). C-130 aircrew coordination behavior: A validation study.

Silverman, D. R. (1995). *MH-53J simulator annual refresher training: Observations and recommendations.* (UDR-TR-95-25). Dayton, OH: University of Dayton Research Inst.

Spiker, V. A., Tourville, S. J., Silverman, D. R., & Nullmeyer, R. T. (in press). *Team performance during combat mission training: A conceptual model and measurement framework.* Mesa, AZ: Human Resources Directorate Aircrew Training Research Division.

Thornton, R. C., Kaempf, G. L., Zeller, J. L., & McAnulty, D. M. 1992). *An evaluation of crew coordination and performance during a simulated UH-60 helicopter mission* (ARI-RN-92-63). Fort Rucker, AL: U. S. Army Research Institute for the Behavioral and Social Sciences.

Tourville, S. J., Spiker, V. A., Silverman, D. R., & Nullmeyer, R. T. (in press). An assessment methodology for team coordination in combat mission training. In *Proceedings of the 18th Interservice/Industry Training Systems and Education Conference.* Orlando, FL.

Wilhelm, J., Roithmayr, P., & Helmreich, R. L. (1992). *An Update to the On-Going Evaluation of the Navy Aircrew Coordination Training Seminar* [Overhead Slides]. Austin, TX.

Zeller, J. L. (1992). *Aircrew coordination evaluation of the AH-64 Instructor Pilot Course* (WP ARIARDA/ASI 92-01). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.