TEAM COORDINATION AND PERFORMANCE DURING COMBAT MISSION TRAINING

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

ABSTRACT

Combat mission training and crew coordination are both integral parts of aircrew training. The research presented here demonstrates a direct empirical link between crew coordination and crew mission performance. Eleven Air Force Special Operations Command (AFSOC) MC-130P aircrew, composed of two Pilots, two Navigators, one Flight Engineer (FE), and one Communication Systems Operator (CSO), were observed during the preparation and execution of a highly complex combat mission scenario in the MC-130P Weapon System Trainer (WST). Five subprocesses of crew coordination were previously identified (Time Management (TM), Function Allocation (FA), Tactics Employment (TE), Situation Awareness (SA), and Command, Control, and Communications (C3)). These subprocesses, along with several mission performance variables (e.g., chart preparation, briefing quality, mission phase performance, etc.), were observed and rated across mission preparation and four phases of execution (Low-Level (LL), Air Refueling (AR), Air Drop (AD), and Infil/Exfil (I/E)). The results demonstrate: (1) a strong positive overall process-performance correlation ($r = .86$); (2) differential impacts of overall process on phase-specific mission performance; (3) differential impacts of mission phase process on overall performance; and (4) differential impacts of particular subprocesses on mission phase performance. We conclude with a discussion of the training implications of our results and observations.

ABOUT THE AUTHORS

Denise R. Silverman is a research psychologist with Hughes Training, Inc. - Training Operations, at Kirtland AFB, New Mexico. She has conducted research on simulator effectiveness, simulator sickness, mental models, CRM, virtual reality, usability, and combat mission training. She has a Ph.D. in Experimental Psychology from the University of New Mexico.

V. Alan Spiker is a principal scientist at Anacapa Sciences, Inc., Santa Barbara, CA. He is responsible for human factors research associated with advanced technology and has conducted field studies on mission planning and command and control systems for all branches of service. He has a Ph.D. in Experimental Psychology from the University of New Mexico and has been with Anacapa Sciences since 1982.

Steven J. Tourville is a senior training analyst and operations manager with Hughes Training, Inc. - Training Operations. He is experienced as a USAF Joint Special Operations mission planner, an instructor navigator, and a training manager in the MC-130P Aircrew Training System. He is responsible for managing field research in the areas of aircrew team performance and training evaluation for combat mission training and mission rehearsal. He is currently an Ed.D. candidate in Adult Education at Nova Southeastern University.

Robert T. Nullmeyer is a research psychologist with the Aircrew Training Research Division of the Armstrong Laboratory at Mesa, AZ. He has conducted research on training system design and evaluation, simulator effectiveness for training and rehearsal, and combat training requirements. He earned a Ph.D. in Experimental Psychology from the University of New Mexico. He joined the Aircrew Training Division of the Air Force Human Resources Laboratory in 1981.

INTRODUCTION

This report describes the results of a study that examined the relationship between the effectiveness of team coordination and tactical performance during a simulated mission that was conducted as a natural phase of MC-130P Special Operations Forces (SOF) combat mis-sion training (CMT). The work was performed at the 58th Training Support Squadron (58 TRSS), Kirtland Air Force Base (KAFB) where Hughes Training, Inc. (HTI) manages an operating loc-ation for the Armstrong Laboratory Aircrew Training Research Division (AL/HRA). A major goal of the HTI research program at KAFB involves the identification, measurement, valid-ation, and eventual reinforcement of aircrew behaviors associated with combat mission readiness and effective mission performance.

Combat Mission Training and Tactical Team Resource Management (T^2RM)

Crew 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 which most differentiate effective crews from weaker ones involve leadership, decision making, and resource management, thereby establishing the need for training "softer" as opposed to more technically-oriented skills.

Since then, many CRM training programs have been developed and various conceptions of the appropriate content of CRM training proposed. Traditionally, CRM has included training in such areas as leadership, communication, decision making, workload management, and situation awareness (Gregorich & Wilhelm, 1993). In designing the present study, we expanded the traditional view of CRM to include subprocesses that are perhaps more applicable to complex, turbulent CMT environments. This view was substantiated during pilot tests, observations, and interviews conducted with SOF subject matter experts (SMEs) (Spiker, Tourville, Silverman, & Nullmeyer, 1996). On the basis of these front-end analyses, we have proposed the concept of "tactical team resource management" or T^2RM as more appropriate for CMT than CRM. T^2RM embraces three elements that distinguish it from CRM:

One, the content of T^2RM has a strong tactical component in which combat skills, rather than attitudes or interpersonal relationships, are emphasized. The conduct of a successful SOF operation places heavy demands on the tactical skills of all crewmembers, where information is uncertain, decisions are split-second, and timing is all-important. Aircraft must be operated at low altitudes, under adverse weather, in the midst of relocatable threats, and with no margin for error (Spiker & Campbell, 1993). Compre-hensive T^2RM measurement ensures that reliable indices are obtained in such tactical areas as navigation accuracy, threat avoidance, meeting control times, and accomplishing critical mission events.

Two, T^2RM emphasizes the behavioral proc-esses of the entire combat team. As used here, a team is composed of those players whose information, actions, and decisions impact the aircrew's mission in some way (Andrews, Bell, & Nullmeyer, 1995). Besides the aircrew, this includes such mission participants as intel-ligence, planners, weather, airborne command, control, and communications, air traffic control, maintenance, as well as the ground "customers" (e.g., Seals) being supported by the aircrew. As such, comprehensive T^2RM measurement ensures that "data hooks" are in place so that key interactions between the aircrew and the extended team are captured, as well as those within the aircrew itself.

Three, T^2RM emphasizes the management of multiple and diverse resources in a dynamic environment throughout the entire course of the mission, *including planning*. A SOF team must ensure that required mission events are per-formed within demanding temporal constraints and that duties are efficiently distributed among a limited cadre of people. SOF teams must leave a minimal "footprint" when performing mission events—people and aircraft are kept to a minimum. Thus, effective T^2RM entails using team members to their fullest during mission preparation and execution (Spiker & Nullmeyer, 1995). Consequently, T^2RM measurement requires data collection throughout *all* mission phases.

T^2RM Measurement Model

Above, several references were made to the measurement requirements for capturing discriminating T^2RM behaviors among crew-

members and aircrews. These include using multiple "data hooks" to capture behaviors and interactions between crewmembers and other team players (e.g., Inteligence), monitoring tactically relevant behaviors, and measuring behaviors across the entire range of mission events. These requirements are embodied in the T^2RM Measurement Model depicted in Figure 1. This model represents a systematic and comprehensive framework for T^2RM data collection in a CMT environment, and estab-lishes the methodological framework for our work.

The concepts and arrows in Figure 1 flow from left to right, reflecting an implicit timeline of SOF MC-130P Annual Refresher Training (ART) activities. The white boxes within the interior of the figure represent ART elements that must be measured during the course of training in order to establish links between T^2RM and crew mission performance. The boxes on the periphery indicate the instruments that were used to collect attitude, performance, process, and outcome data on crewmembers and aircrews.

In this paper we focus on the T^2RM and Team Mission Performance/Mission Outcome mod-ules, where information was collected primarily using the Team-Mission Observation Tool (T-MOT) and the Team-Mission Performance Tool (T-MPT), respectively.

The independent obser-vations recorded on these tools enable key relationships between crew process (T^2RM) and team performance to be drawn.

The acronyms for the five T^2RM subprocesses we examined are depicted in the T^2RM module (see Table 1 for their formal definitions). They are represented as cubes because they were assessed across five mission phases and across each crew position. These five subprocesses were selected for study based on their: judged relevance to the AFSOC mission environment, appropriateness to the high levels of experience of MC-130P aircrews reporting to ART, applicability to CMT, and amenability to measurement by outside observers (Spiker et al., 1996). As well, we attempted to identify subprocesses that made contact with the CRM dimensions previously studied by other researchers.

Since we know that these five areas do not encompass the entire domain of what would properly be considered T^2RM, we denote the presence of other tactically relevant processes by the empty gray-shaded cube within the T^2RM module. In addition, the Other CRM Processes module is depicted in the gray-shaded box beneath the T^2RM module. These include such
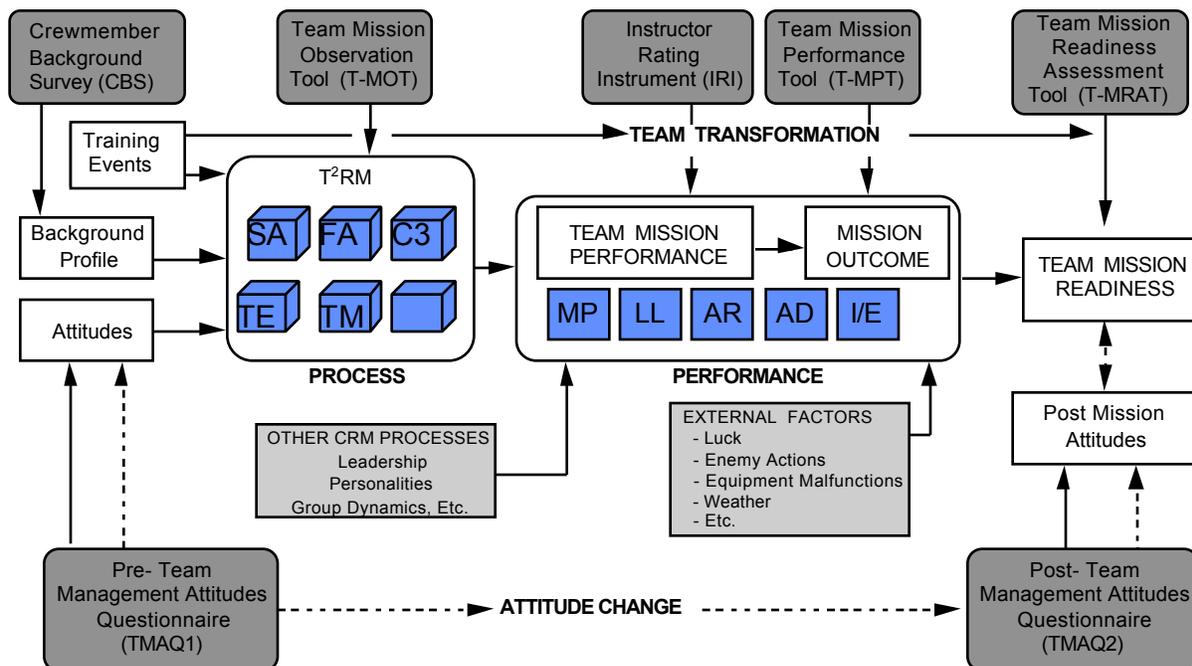

Figure 1. T^2RM Measurement Model (Spiker et al., 1996).

Table 1. Definition of Five T^{2}RM Subprocesses.

T^{2}RM Subprocess	Definition
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.
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.
Situation Awareness (SA)	The maintenance of an accurate mental picture of mission events and objectives as they unfold over time and space.
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.
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.

traditional process variables as group cohesiveness, personality, group dynamics, and leadership. While these factors may have some influence on team mission performance, they were not measured in our research.

The output of the T^{2}RM module feeds into the Performance module. By team mission perfor-mance, we mean those indices that directly result from the successful (or failed) execution of important T^{2}RM processes. In our research, team mission performance is reflected in such indices as the quality of the mission briefings, completeness of the navigation charts, and instructor-supplied ratings of how well the team as a whole and individual crewmembers executed each mission phase. The boxes above the Team Mission Performance module list the two tools (Instructor Rating Instrument or IRI and the T-MPT) we used to collect this information.

Team Mission Performance and Mission Outcome both reside within the larger Performance module to indicate their strong relationship and that some of our team perfor-mance indices include data that would normally measure mission outcome. Mission outcome is reflected in indices that indicate whether the team accomplished its stated objectives. When performed in a weapon systems trainer (WST), some of these tasks can be recorded by computer. Outcomes include airdrop accuracy, performing airlands within prescribed time windows, on-ground time for transloads, and minimizing (or avoiding) exposure to threats. These were used as behavioral anchors in the T-MPT to assess performance.

T^{2}RM and mission performance were assessed for each crew position present for ART across each mission phase, as indicated by the shaded boxes within the Performance

module. For MC-130P ART, this includes six crew positions and five mission phases. The standard MC-130P aircrew consists of seven crewmembers. Five members perform their primary flight duties in the cockpit: Aircraft Commander (AC), Co-Pilot (CP), Left and Right Navigators (LN, RN), and Flight Engineer (FE). The Communication Systems Operator (CSO) and Load Master (LM) crew stations are located in the rear cabin compartment. The LM, however, is omitted from further consideration, as this crewmember's ART is not conducted with the other crewmembers at the 58 TRSS. The five MC-130P mission phases we observed during ART were Mission Preparation (MP), Low-Level (LL), Air Drop (AD), Air Refueling (AR), and Infil/Exfil (I/E). Within each phase, there are multiple opportunities for the crew to engage in behaviors that fall within each of the five T^{2}RM subprocesses.

In sum, our T^{2}RM measurement model embodies a number of features that are critical for establishing direct links between team coordination processes and mission perfor-mance and for determining some of the underlying factors that potentially influence this relationship. Key features include: (a) providing multiple measures of team performance in addition to mission outcome; (b) assessing a series of tactically relevant coordination processes; (c) measuring team coordination process and performance throughout the timeline of training events and across crew positions; (d) obtaining baseline measures of individual crew acceptance of CRM training and principles; and importantly, (e) emphasizing independent assessments of team coordination process and mission performance.

T^{2}RM and Mission Performance

Though it is commonly assumed that effective aircrew coordination results in superior mission performance, very few studies have demon-strated an empirical link using tactically realistic training scenarios. Povenmire, Rockway, Bunecke, and Patton's (1989) study of B-52 crew coordination represents one of the strongest demonstrations of a direct rela-tionship between crew coordination processes and mission performance. They used a number of the methodological features depicted in our T^2RM measurement model.

Povenmire et al. (1989) 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. Highly trained CRM evaluators assessed crew coordination and mission performance, with separate sets of raters used for each measure. Three measures of mission performance were taken: bombing accuracy, threat avoidance, and technical skill. The primary analysis revealed a strong, positive correlation (r = .84) between crew coordination and mission performance. Povenmire et al. (1989) also compared the experts' ratings of mission performance with the individual mission outcome factors, revealing that the experts primarily used bombing accuracy to make their overall judgments of mission performance. Similarly, overall crew coordination was compared with individual coordination dimen-sions. This analysis showed that raters based their judgments of crew coordination on four dimensions—practicing inquiry and advocacy, avoiding distractions, distributing workload, and resolving conflicts.

Primary Research Hypotheses

Despite the simplicity of its design and data analysis strategy, the Povenmire et al. (1989) study stands as one of the clearest demonstrations of the relationship between aircrew coordination processes and mission performance. Indeed, the elegance of their design provides unusually clear insights regarding the crew coordination processes that best predict mission performance. We used many of Povenmire et al.'s (1989) method-ological features in our empirical study of T^2RM. The two main hypotheses we tested were:

(1) Team coordination processes should be strongly and positively related to mission performance. Crews who exhibit superior coordination behaviors should perform better during mission execution than those who do not. If found, this would extend the Povenmire et al. (1989) finding with B-52 crews to another weapon system. As described in Spiker et al. (1996), the process-performance relationship is the fundamental tenet of the proposed research. A significant finding "gives us permission" to probe the data further, to test the second hypothesis as well as those explored in Spiker, Silverman, Tourville, and Nullmeyer (in press) and Silverman, Spiker, Tourville, and Nullmeyer (in press).

(2) The various T^2RM subprocesses should show differential relationships to per-formance across mission phases. Although we expect an overall process-performance relationship, we do not expect every T^2RM process to be significantly correlated with every phase of mission performance. We do not believe T^2RM to be single entity. Rather, for the performance in a given mission phase, only one or two T^2RM subprocesses may be statistically significant predictors. For example, the FA subprocess may have a significant impact on those phases of performance (e.g., LL) where allocation of crew duties can substantially reduce overall team workload. Such selectivity is a natural outcome, given that different phases will require different types of crew involvement, interaction, and decision making.

METHOD

Participants

Thirteen MC-130P SOF aircrews were observed during their week-long visit to KAFB for ART. Two crews were eliminated from the analyses because they were unable to fly the mission due to simulator malfunctions. Thus, 11 crews (67 crewmembers) were included in the final analysis. The participants were experienced SOF personnel; average reported flying hours was 3056, 1286 of those in the MC-130P. The typical make-up of an MC-130P crew for ART is six members. In this study, crew size varied from 5-7 as two crews trained with an extra CSO for the observed mission and one crew did not have a CSO.

Data Collection Instruments

Six different data collection tools were used, but only two are relevant to this paper:

Team-Mission Observation Tool (T-MOT).
The T-MOT was an integral part of our total
assessment strategy. This instrument was
designed to aid in recording specific
individual and team coordination behaviors
that fell into distinct T^2RM subprocesses, and
which occur during each mission phase.
Measurement was accomplished using five-
point Likert scales and SME observations of
critical coordination behaviors during the
complex CMT scenario.

T-MOT data were collected using over-the-
shoulder structured observations. While there
was a "script" to help our SME structure his
observations, he was also free to record "by
exception" activities, attitudes, behaviors, and
cognitions that seemed unusually strong or
weak. The T-MOT supported the recording
and analysis of both individual crewmember
and crew behaviors within the five T^2RM
subprocesses across the five mission phases.
This qualitative information supplemented
quantitative analyses performed on the rating
data that were also obtained with this
instrument.

The five-point rating scale (1=lowest to
5=highest) provided quantitative assessments
across T^2RM subprocesses, crewmember
positions, and mission phases. This
technique proved to be a fairly efficient way
to generate a large amount of data within a
well-defined structure in which the observer
used rules to assign scale ratings to specific
attributes of team coordination. A complete
description of the T-MOT can be found in
Tourville, Spiker, Silverman, and Nullmeyer
(1996).

Team-Mission Performance Tool (T-MPT).
The T-MPT was designed to aid in evaluating
crew mission performance demonstrated in
each mission phase. This instrument
provided several behaviorally anchored
ratings scales (BARS) for a second researcher
to rate the quality of individual- and team-
generated mission products developed during
MP and demonstrated crew performance
across mission phases.

Procedure

The process data were recorded on the T-
MOT by a highly trained, former MC-130P
navigator. During MP, this individual served
as a *participant-observer (P/O)*. This
relationship was deemed appropriate as our
SME possessed a comprehensive

understanding of current SOF doctrine and
had extensive flight, training, and simulation
experience in SOF operations through
training exercises and operational missions.
In addition, the P/O approach has been
successfully used in other team training
research realms (e.g., Oser, Dwyer, & Fowlkes,
1996). The P/O's primary responsibility was to
perform data collection tasks (observer) using
the T-MOT. Secondarily, he was a
"controller" (participant) which involved role-
playing as liaison officer with outside
(simulated) agencies and providing mission
debrief support.

During mission execution, the SME-
researcher observed and monitored crews
from an intercom station located outside the
MC-130P WST, situated in front of four
instructor-operator screens that repeated the
instructor inputs from inside the MC-130P
WST. From this vantage point, the SME-
researcher was able to make informed
observations and assessments based on
personal knowledge and expertise about the
mission and crew coordination, and
monitoring instructor inputs.

The second researcher observed crew
performance keeping in mind the criteria
established in the T-MPT. Her notes and
observations captured such items as: the
number of briefings each crew gave, the
contents of the briefings, who performed the
briefings, the number of charts created, etc.
All of these notes facilitated post-mission
completion of the T-MPT. Similar types of
notes were taken during mission execution
(e.g., time of AD). This researcher was
purposely located away from the process
researcher during mission execution. From
her location, she monitored the crew
communications, flight path, and threat
laydown. It was also from this location that
various Instructor Operator Station (IOS)
pages were selected and printed out, to aid in
later completion of the T-MPT. At the
conclusion of the simulator mission, the
researcher collected all products (flight plans,
charts, checklists, etc.) that the crew created
for the mission. These were collectively
evaluated using the T-MPT. For example,
AD performance was evaluated using the T-
MPT BARS for AD, where the actual
performance was assessed using the IOS
ground track map page, the prepared chart,
and the recorded drop time.

RESULTS

The analyses that we report in this subsection involve Pearson product moment correlations between various pairs of *crew-level* rating data associated with T^2RM and combat mission performance. These correlations are summarized in Table 2.

Data Structure

Before discussing the results, we first describe some conventions for interpreting the contents of this rather complex summary table. The first column indicates the types of analyses con-tained in the subsequent cells of that row. Column 2 refers to correlations between pairs of performance and process ratings that reflect the whole mission; the remaining five columns involve ratings of behaviors performed within specific mission phases.

The rows in the table correspond to categories of correlations that cross mission phases. The top row depicts correlations between our most global process rating (the overall process rating for the entire mission) and the five mission phase performance ratings. The second row summarizes correlations between our most global performance rating, performance sum (PerfSum), and the overall process ratings generated for each mission phase. The third row summarizes correlations between the phase-specific overall process ratings and phase-specific performance ratings. The remaining five rows reflect correlations between performance ratings and each of the five subprocess ratings within each of the six mission categories. It should be noted that the first column contains correlations between PerfSum, which was computed as a sum of the five mission phase ratings, and a

summary rating for each subprocess that takes all mission phases into account. We used a sum of the five mission-phase performance ratings because the skills required in each phase are so different that we did not believe that an overall rating would be as valid.

Within this structure, the most basic question we ask is whether there is a statistically significant relationship between mission performance (PerfSum) and overall crew T^2RM across the entire mission. The first row, first column depicts the correlation between the overall rating of crew coordination for the whole mission and the rating of crew performance for the total mission (PerfSum). Outlined in bold, this whole-mission T^2RM by PerfSum correlation is the single most important analysis in the table, as it provides our most global, and hopefully most robust, assessment of the T^2RM-CMT performance relationship.

To further explore the relationships between process and performance, we looked within the five primary mission phases. Our data structure allowed three ways to view phase-specific process and performance relationships: expand performance ratings to reflect individual mission phases and correlate these with the whole mission T^2RM rating (Row 1), expand process ratings and correlate with PerfSum (Row 2), or expand both (Row 3).

The five entries on the right hand side of the top row depict correlations between the whole-mission, overall process rating, and the five performance ratings for specific mission elements. For this row, the five column headings which contain the analytic results— MP, LL, AD, AR, and I/E—reflect the origin of the performance ratings.

Table 2. Correlations Among Ratings of T^2RM Processes and Combat Mission Performance.

Process/Performance	Whole Mission	MP	LL	AD	AR	I/E
Overall Process x Performance	.86*	.78*	.74*	.39	.52	.41
Phase-Specific Process x Performance Sum		.78*	.57	.79*	.75*	.77*
Phase-Specific Process x Phase-Specific Performance		.68	.46	.54	.70	.67
SA Subprocess x Performance	.76*	.48	.47	.59	.28	.65
TE Subprocess x Performance	.78*	.27	.06	.54	.81*	.55
TM Subprocess x Performance	.83*	.41	.36	.66	.51	.64
FA Subprocess x Performance	.75*	.22	.61	.55	.55	.60
C3 Subprocess x Performance	.08	.14	.09	.32	.37	.30

$*p_{EW}<.05$, $p_{NOM}<.001$ Bonferroni adjustment assuming 46 tests; critical r = .74

The second row in Table 2 depicts the correlations between phase-specific overall process ratings and PerfSum. The five column headings which contain the analytic results now refer to the phase in which the overall process measure was taken, not performance. The first data cell in that row has been shaded out since its combination is identical to the corresponding cell in the first row (i.e., overall T^2RM and overall performance).

The third row in Table 2 follows a similar logic to that in the first and second rows. This time, however, the values refer to correlations between phase-specific process ratings and phase-specific performance ratings. Thus, the five mission phase column headings now contain ratings in which **both** process and performance are phase-specific.

The correlations in the lower part of the table explore the impact of T^2RM further, by consid-ering the relationships between the various T^2RM **subprocesses** and the corresponding performance rating. The five lightly shaded cells in the second column depict the correlations between each of the overall subprocess ratings (i.e., assessed across mission phases) with total mission performance, or PerfSum. The remaining 25 cells contain the most detailed look at the process-performance relationship. They depict the correlations between each subprocess—within a each mission phase—with the corresponding performance rating for that mission phase.

Statistical Testing Considerations

Since all the tests reported in this section use crew as the unit of analysis, our total N (11) and resulting nine degrees-of-freedom (for t-tests, df = N-2 = 9) would seem to yield insufficient statistical power to establish a strong process-performance relationship. However, the 11 crews in our sample constitute 26% of the population of 42 SOF MC-130P aircrews that complete ART each year. Because we sampled a sizable proportion of the population, we are able to reduce our estimated variance of the sample mean by using a finite-population correction coefficient (Winkler & Hays, 1975). The correction coefficient decreases the observed sample variance by the square root of (N-1)/N-n, where N is the population size and n is the sample size. In our case, the reported t-values have been increased by 20%, reflecting a 1.2 finite-population coefficient multiplier.

Due to the exploratory nature of many of our research questions, and the need to perform a large number of statistical tests, we used a Bonferroni adjustment as a way to keep our overall, or experimentwise (EW), alpha level from exceeding the desired (nominal) level. The Bonferroni technique is a fairly conservative, though effective, way to avoid inflating the alpha level (and hence the likelihood of a Type I error) by "snooping" through one's data to locate the largest effect (Harris, 1994). The adjustment is made by dividing the desired EW alpha level by the number of tests that are performed in a given cycle of testing.

As shown at the bottom of Table 2, we have employed a fairly conservative nominal alpha level of .001 to control for the fact that we are performing 46 comparisons. Since all correl-ations are tested against a null hypothesis of 0, this corresponds to a critical t-value of 4.19 (df = 9, one-tailed). Given the finite population correction described above, this means that our reported correlations have to reach a criterion of at least .74 to be statistically significant.

Findings

Turning to the correlations in Table 2, we find that our primary hypothesis—a strong, positive relationship between overall T^2RM process and overall mission performance—is supported by the data. This is evident from the large correlation coefficient, .86, that appears in the upper left-hand cell of the table (t = 6.143, df = 9, p < .05, two-tailed). Figure 2 depicts the scatterplot between crew performance sum on the x-axis and T^2RM rating on the y-axis. As is evident from the figure, the poorest performing crews did indeed have the lowest overall T^2RM process rating whereas the highest rated crews received the largest crew performance sums.
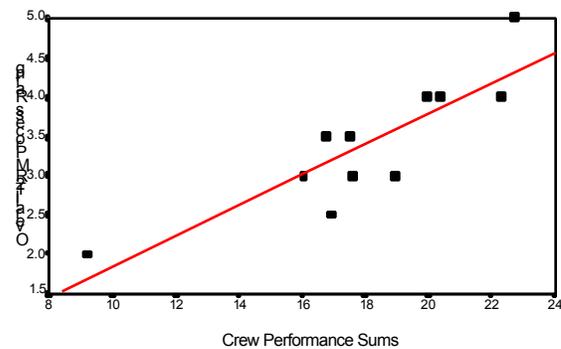

Figure 2. Scatterplot of Overall T^2RM Process Ratings and Mission Performance Sums.

Having established the fundamental relationship between T^2RM and mission performance, the next step entailed

pinpointing the subprocesses and/or phases in which the relationship is the strongest. As a first step, note that the five correlations in the top row of Table 2 depict the correlations between overall T^2RM and the performance rating within each mission phase. As denoted by the asterisks, we see that the quality of T^2RM is significantly related to only two of the mission phases, MP and LL. Though trending positive, performance in the other three phases was not significantly related to T^2RM process. This finding shows that while the quality of the T^2RM is important for performance, it is not uniformly so, but is instead somewhat localized. The large influence of these phases may be due to their relative durations compared to the AD, AR, and I/E phases. That is, both T^2RM subprocess and team mission performance ratings were based on sampling behavior throughout extended periods of time (3 hours for MP, blocks of minutes throughout mission execution for LL), thus perhaps providing more representative and stable performance measures.

Examination of the correlations in the next two rows of Table 2 shows the influence that phase-specific T^2RM exerts on overall performance (second row) and phase-specific mission performance (third row). As seen in the second row, four of the phase-specific T^2RM measures significantly predict overall mission perfor-mance. The only phase that failed to achieve significance was LL. This finding suggests that the influence of T^2RM on performance is not localized in a single phase, but rather, is present throughout the entire mission scenario.

The gray-shaded cells in Table 2 depict the correlations that gauge the strength of the linear relationship between each of the five T^2RM subprocesses and overall mission performance (i.e., PerfSum). Four of the sub-processes were statistically related to mission performance, with only C3 failing to achieve significance. The pattern of results in these cells suggests that our decomposition of T^2RM into its constituent subprocesses was successful in identifying those factors having a strong impact on performance.

The lower right quadrant in Table 2 (outlined in bold) breaks down the analysis still further, as it comprises a 5 x 5 matrix of correlations between each phase-specific T^2RM subprocess ratings and the corresponding phase-specific crew performance ratings.

While all of the correlations were positive, and some quite large, only the correlation between the rating of TE and crew performance during the AR phase proved significant (t = 4.97, df = 9, p < .001, one-tailed). Though our statistical tests cannot substantiate this, there is some evidence that the T^2RM subprocesses exhibit differential effects across mission phases (Hypothesis 2). Additional evidence for this differential sensitivity comes from a series of multiple regression analyses (MRAs) that we performed across the five phases. In each case, the analysis showed that beyond the first or second-highest correlated T^2RM subprocess (e.g., SA for MP), little variance is accounted for by adding other subprocesses into the MRA equations.

This trend is also evident in the pattern of correlations themselves. Examining each column, we see that the T^2RM subprocess most strongly associated with the MP phase is SA, with a smaller contribution by TM. The main T^2RM subprocess associated with LL is FA, with some impact of SA. The T^2RM subprocess most strongly associated with AD performance is TM. Besides TE, other T^2RM subprocesses influencing AR performance were TM and FA. Finally, I/E performance was associated with fairly large (though nonsignificant) correlations by all of the subprocesses, save C3. Indeed, as can be seen in the bottom row of Table 2, C3 did not have any correlations above .40 in any of the mission phases. This is consistent with the low correlation that the C3 subprocess overall rating exhibited with overall mission performance. These results suggest that our reported subprocess measures of T^2RM have the potential to be fairly sensitive barometers of team mission performance.

CONCLUSIONS

The most salient finding in this study was the dramatic replication of the Povenmire et al. (1989) observation that crew coordination process is a significant predictor of tactical mission performance. In our study, this rela-tionship achieved a correlation of .86. While there have been descriptions of the relationship between crew coordination and performance in the context of accident investigations and safety violations (e.g., Predmore, 1991), direct empirical support has been scarce, especially in tactical environments.

The significant correlation between T^2RM process and mission performance then afforded us the opportunity to explore the relationships of particular subprocesses and mission phases, to overall performance and T^2RM. We have also documented the specific behaviors that underlie these

relationships (Spiker et al., in press; Silverman et al., in press). These detailed analyses provide us with the critical information required to: (a) provide CRM-CMT training content input, (b) improve our data collection instruments, and (c) further our research efforts.

One of the primary CRM training implications of these findings is the need to introduce more weapon system specific content (e.g., T^2RM subprocesses) into CRM training. The AFSOC approach to CRM does not presently address aircraft- or mission-specific requirements.

Several lines of evidence support this conclusion. First, our high correlations were possible only because we found a wide range of CRM behaviors and a wide range of mission performance levels across mission qualified crews. Second, instructor comments suggest that the current lack of solid *tactically relevant* CRM skills, like those observed and rated as T^2RM subprocesses, make CRM training difficult. The behaviors (Spiker et al., in press) associated with each of the subprocesses and crew positions could be incorporated into future CRM training. Third, forward-looking practitioners of CRM training, who are held up as leaders in this business, have begun to discard generic CRM training practices and now incorporate notions of operational relevance and tailored training into their CRM programs.

REFERENCES

Andrews, D. H., Bell, H. H., & Nullmeyer, R. T. (1995). Aircrew team training: A tutorial. In *Proceedings for The 17th Interservice/Industry Training Systems and Education Conference.* Albuquerque, NM.

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.

Harris, R. J. (1994). *ANOVA: An analysis of variance primer.* Itasca, IL: F.E. Peacock Publishers, Inc.

Oser, R., Dwyer, D. J., & Fowlkes, J. E. (1995). Team performance in multi-service distributed interactive simulation exercises: initial results. In *Proceedings of the 17th Interservice/ Industry Training Systems and Education Conference.* Albuquerque, NM.

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 cooridination in United 811 and United 232. In *Proceedings of the Sixth International Symposium on Avition 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.

Silverman, D. R., Spiker, V. A., Tourville, S. J., & Nullmeyer, R. T. (in press). *Impact of crew position process on team performance during combat mission training.* (AL/HRA-TR-xxxx). Mesa, AZ: Armstrong Laboratory, Aircrew Training Research Division.

Spiker, V. A. & Campbell, J. L. (1993). *Functionial analysis of Combat Talon I mission planning: Opportunities for integration with advanced mission rehearsal technology* (AL/HRA-TR-1993-0058). Mesa, AZ: Armstrong Laboratory, Aircrew Training Research Division.

Spiker, V. A. & Nullmeyer, R. T. (1995, June). *Functional analysis of the MH-53J Pave Low mission preparation process.* (AL/HR-TR-1994-0171). Mesa, AZ: Aircrew Training Research Division, Armstrong Laboratory.

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

Spiker, V. A., Tourville, S. J., Silverman, D. R., & Nullmeyer, R. T. (in press). *Effects of tactical team resource management (T^2RM)* (AL/HR-TR-1997-xxxx). Mesa, AZ: Aircrew Training Research Division.

Tourville, S. J., Spiker, V. A., Silverman, D. R., & Nullmeyer, R. T. (1996). An assesment methodology for team coordination in combat mission training. In *Proceedings of the 18th Interservice/Industry Training Systems and Education Conference.* Orlando, FL.

Winkler, R. L. & Hays, W. L. (1975). *Statistics: Probability, inference, and decision.* New York: Holt, Rinehart, and Winston.