GUIDELINES FOR FACILITATING SHIPBOARD TEAM TRAINING

Sandra S. Bailey   Joan Hall Johnston   Kimberly Smith-Jentsch   Gregory Gonos   Janis A. Cannon-Bowers

Naval Air Warfare Center Training Systems Division
Orlando, FL 32826

Tactical decision-making teams in the modern warfare environment are faced with situations characterized by rapidly unfolding events, multiple plausible hypotheses, high information ambiguity, severe time pressure, and severe consequences for errors. In order to adapt to these stressors, team members must learn to coordinate their actions so that they can gather, process, integrate, and communicate information in a timely and effective manner. A key factor toward ensuring a team's success in a cognitively complex and stressful task environment is training that incorporates explanation, demonstration, practice, dialogue and feedback. However, evaluating teamwork skills, and providing meaningful performance feedback is a complex and demanding task for combat systems' trainers. In support of facilitating real-time shipboard team training assessment and feedback, the Naval Air Warfare Center Training Systems Division developed the On-line Anti-Air Warfare Team Observation measure (ATOM) and Debriefing Guide. These performance measurement tools facilitate the capture of on-line performance measures, and provide diagnostic feedback to a team. Both instruments were tested in conjunction with the Battle Force Tactical Training developmental test on-board an AEGIS capable ship in October 1994. The On-line ATOM was used to assess team processes with respect to each scenario, and shipboard trainers used the Debriefing Guide to provide feedback to the team within 15 minutes following each scenario. Trends during the DT-IIA training period showed an improvement in the team's teamwork skills. Therefore, the purpose of this paper is to provide guidelines for facilitating shipboard team training. Guidelines will include (a) designing trigger event-based training scenarios, (b) developing measures of team performance, (c) designing a performance feedback system, and (d) developing training for assessing and providing feedback to facilitate teamwork skills.

About the Authors

Dr. Sandra S. Bailey is a research psychologist at the Naval Air Warfare Center Training Systems Division. She is responsible for conducting behavioral research to develop training and simulation principles for tactical decision-making under stress. She received her Ph.D. in Human Factors Psychology from the University of Central Florida.

Dr. Joan Hall Johnston is a research psychologist at the Naval Air Warfare Center Training Systems Division. She is responsible for conducting behavioral research to develop training and simulation principles for the Tactical Decision-making Under Stress exploratory research program. She received her Ph.D. in Industrial/Organizational Psychology from the University of South Florida.

Dr. Kimberly Smith-Jentsch is a research psychologist at the Naval Air Warfare Center Training Systems Division. She is responsible for conducting behavioral research to develop training and simulation principles for the Tactical Decision-making Under Stress exploratory research program. She received her Ph.D. in Industrial/Organizational Psychology from the University of South Florida..

Greg Gonos is a research psychologist at the Naval Air Warfare Center Training Systems Division. He is responsible for conducting transfer of training research and development. He received his M.S. in Industrial/Organizational Psychology from the University of Central Florida. He is currently working on doctorate in Human Factors Psychology at the University of Central Florida.

Dr. Janis A. Cannon-Bowers is a senior research psychologist at the Naval Air Warfare Center Training Systems Division. She is co-principle investigator for the Tactical Decision-making Under Stress program, and is responsible for behavioral research concerned with improving individual and team decision-making in the Navy tactical environment. She received her Ph.D. in Industrial/Organizational Psychology from the University of South Florida.

GUIDELINES FOR FACILITATING SHIPBOARD TEAM TRAINING

Sandra S. Bailey   Joan Hall Johnston   Kimberly Smith-Jentsch   Gregory Gonos   Janis A. Cannon-Bowers

Naval Air Warfare Center Training Systems Division
Orlando, FL 32826

INTRODUCTION

Tactical decision-making teams in the modern warfare environment are faced with situations characterized by rapidly unfolding events, multiple plausible hypotheses, high information ambiguity, severe time pressure, and serious consequences for errors (Cannon-Bowers, Salas, & Grossman, 1991). In order to adapt to these situations, team members must learn to coordinate their actions so that they can gather, process, integrate, and communicate information in a timely and effective manner. The accurate diagnosis of performance shortfalls, and the tailoring of subsequent training toward correcting these shortfalls for teams and team members are contingent upon systematic performance assessment from multiple perspectives. Unfortunately, it has been the case in the past that operational systems either ignore performance measurement completely, or treat it in an unsystematic, "piecemeal" fashion. This is particularly true of complex systems where it is difficult to assess performance with a single correct answer, or in situations where there are several individual decision makers who must interact as a team.

Previous research efforts in the area of naturalistic team decision-making (Klein, 1993) and team training (Cannon-Bowers, Tannenbaum, Salas & Volpe, 1995) have shown that teamwork skills are separate and distinct from task work skills, and hence require dedicated training. Expert teams adjust their responses in accordance to the task demands (stress), and employ implicit coordination strategies (without the need to communicate) by drawing upon common knowledge bases (Salas, Cannon-Bowers & Johnston, in press). Team training research programs have generated guidelines and strategies for team training. Dwyer (1992) developed an index of team performance that is based on observable indicators of effective and ineffective behaviors across critical functions of an AAW team. Cannon-Bowers, Salas, and Grossman (1991) suggested methods for enhancing tactical decision-making of Anti-Air Warfare (AAW) Combat Information Center (CIC) teams under stress.

A key factor in ensuring a team's success in a cognitively complex and stressful task environment is training that incorporates explanation, demonstration, practice, dialogue, and feedback. However, evaluating team member task work and teamwork skills, and providing meaningful performance feedback "real time" are complex and demanding tasks for combat systems trainers. In support of the developmental test (DT-IIA) Training Effectiveness Validation (TEV) of the Battle Force Tactical Trainer (BFTT), the Tactical Decision-making Under Stress (TADMUS) program at the Naval Air Warfare Center Training Systems Division (NAWCTSD) developed the On-line Anti-Air Warfare Team Observation Measurement (ATOM) tool and Debriefing Guide. During DT-IIA TEV, these instruments were tested on-board an AEGIS capable ship in October 1994. The purpose of this paper is to provide guidelines for facilitating shipboard team training based on the results of the DT-IIA TEV. Guidelines will include (a) designing trigger event-based training scenarios, (b) developing measures of team performance, (c) designing a performance feedback system, and (d) developing training for assessing and providing feedback to facilitate teamwork skills. In additions, future implications.

APPROACH

One of the requirements for the BFTT DT-IIA TEV test was to develop a team performance measurement methodology that could be used for debriefing an Anti-Air Warfare Combat Information Center ship team. During the conduct of the BFTT DT-IIA TEV test, a methodology was developed and tested by the TADMUS program. The methodology would allow shipboard trainers to: (a) assess a team's performance on-line, (b) prepare a debrief based on these assessments within 15 minutes and (c) provide feedback to the team that would enhance teamwork in subsequent scenarios.

In order to achieve this overall objective, a number of subgoals were accomplished. First, training scenarios were developed, and target team behaviors were identified. Second, a performance measurement instrument was developed to aid raters in observing, recording, and making assessments of targeted behaviors on-line. Third, a methodology was developed for preparing, and delivering performance feedback to teams based on these assessments in a manner that facilitated learning and improved performance. The final goal with respect to performance measurement and debrief methodology was to train raters to use the assessment tools and to adapt them to new scenarios. Each subgoal is discussed below.

Scenario Design & Development

Based on other scenario development (Prince, Oser, Salas, & Woodruff, 1993), the findings from the TADMUS research programs have recommended structuring team training simulator scenarios with embedded trigger events that are designed to elicit appropriate tactical decision-making processes and outcomes (Salas, Cannon-Bowers, Johnston, in press; Johnston, Cannon-Bowers & Smith-Jentsch, 1995). Realistic operational conditions and stressors in terms of workload and level of ambiguity can be inserted into the scenarios so that a representative sample of decision-making actions can be observed and measured (Hall, Dwyer, Cannon-Bowers, Salas, & Volpe, 1992).

To evaluate whether team training objectives were achieved, scenario trigger events needed to be developed so that measures of team performance could be evaluated in a standardized, relevant, valid, and reliable manner. During the BFTT DT-IIA TEV, two pairs of scenarios were developed using subject matter experts from the fleet. Each pair consisted of a moderate workload and a high workload version of the same trigger event structure. Workload was manipulated by increasing the number of tracks and extraneous environmental cues.

Each scenario was composed of three main "trigger events." A trigger event is represented by clusters of related scenario stimuli (e.g., the appearance of an air track that the team must correctly identify) that have been demonstrated in laboratory experiments to differentiate between effective and ineffective teams (Entin, Serfaty, & Decker, 1994). Each trigger event imposed similar tactical decision-making demands that an AAW CIC team would usually experience in littoral warfare operations. For each trigger event, individual task behaviors and teamwork behaviors were identified and clearly defined by subject matter experts.

Team Performance Measurement

The context in which decisions are made in a CIC is exceedingly complex. Therefore, a detailed and systematic approach for evaluating team performance is required. Human performance can be depicted in terms of unit of analysis and level of analysis (refer to Table 1) (Canon-Bowers & Salas, in press).

Table 1

Human Performance Measurement Scheme

	INDIVIDUAL	TEAM
P R O C E S S S	• Decision Making Processes • Task Strategy • Information Seeking	• Coordination Behaviors • Communication Flow • Team Strategies
O U T C O M E	• Latency • Accuracy • Timeliness	• Mission Effectiveness • Aggregate Latency & Accuracy • Error Propagation

The unit of analysis distinguishes between individual behaviors and team behaviors. The level of analysis distinguishes between behaviors that represent processes, and behaviors that represent outcomes. Specifically, process measures determine how the task was accomplished, while outcome measures describe what happened as a result of the processes. For example, individual performance processes are represented by information seeking, task strategies, and cognitive decision-making processes, whereas, individual performance outcomes are represented by latency, accuracy, and decision biases. Team performance processes are represented by communication flow, coordination behaviors, and team strategies, whereas, team performance outcomes are represented by aggregates of latency and accuracy. By dividing performance evaluation among the four quadrants, performance measurement strategies can be specified for individual processes, individual outcomes, team processes, and finally team outcomes.

The trigger event-based scenario structure facilitates the development of a performance measurement tool that allows for on-line measurement of team processes. TADMUS findings, and results from other team training research, provided the basis for developing a trigger event-based measurement device known as the ATOM (Anti-Air Warfare Teamwork Observation Measure). Initial laboratory tests with AAW TADMUS scenarios had demonstrated the ATOM's sensitivity to capture the following four

critical team-process measures (Johnston, Cannon-Bowers & Smith-Jentsch, 1995):

1. **Situation Assessment** - communication that promotes team awareness of the surrounding environment, both external and internal to the team.

2. **Communication** - clear and efficient exchange of information using proper terminology, standard procedures for internal and external communications, and use of proper phraseology.

3. **Compensatory Behavior** - monitoring the activities of other team members, taking action to correct team errors, giving and receiving performance feedback in a non-defensive manner, and providing and seeking assistance or backup when needed

4. **Team Leadership** - providing needed guidance to other team members, helping team members focus their activities appropriately and anticipate tasks that should be performed, and providing instruction to other team members to enable them to perform or complete their tasks. Any member can perform the leadership function.

The original ATOM was used as a post-training exercise diagnostic tool. For the BFTT DT-IIA TEV test, the design structure of the ATOM had to be revised to facilitate on-line recording of the raters' observations of the team's performance during a training scenario. The instrument had to be sensitive enough to capture team process measures as they occurred along the scenario event time line. In addition, the format of the collected data had to facilitate the preparation and delivery of a debrief within 15 minutes of the conclusion of the scenario (BFTT DT-IIA Test Plan, 1994).

One of the design considerations was the ease with which an On-line ATOM could be adapted to new scenarios. Using subject matter experts in data collection instrument design, the performance measurement tool was reconstructed and tested. The On-line ATOM was adapted to four training scenarios in four weeks for use during BFTT's DT-IIA TEV.

Debriefing Guide

The BFTT training methodology used during the DT-IIA TEV test incorporated basic principles of learning and effective instruction, and transitioned many of the research findings from the TADMUS Exploratory Research Program to an on-board training exercise. One of the major elements transitioned to the BFTT training methodology was

the method for developing and delivering effective debriefs.

A BFTT TEV test performance criteria required a tool that would allow ship's personnel to prepare a debrief for ship's company within 15 minutes (BFTT DT-IIA Test Plan, 1994). The debrief had to be both prescriptive and diagnostic in nature. TADMUS researchers developed a Debriefing Guide to assist in the preparation and delivery of a debrief on the AAW team's teamwork performance. The utility of the guide was tested during DT-IIA.

At the conclusion of a scenario, the raters held a consensus meeting to discuss the team's performance, and to prepare the debrief. Using the On-line ATOM worksheets, the raters and a designated debriefer discussed each event, and identified effective and ineffective teamwork behaviors that were observed during the training exercise. Data from the On-line ATOM were transferred to the appropriate sections in the Debriefing Guide. The format of the guide facilitated the easy integration of the data in a short period of time. In addition, the raters assessed the team's strengths, and identified areas for improvement for each teamwork dimension. Team goals for each dimension were identified by the raters, and the team's performance on each of team process dimension was rated using an 5-point scale with behavioral anchors (refer to Figure 1). Space was provided for the debriefer to record specific actions taken by the team. This process is repeated for each of the four teamwork dimensions.

While the debrief was being prepared by shipboard trainers, the AAW team assessed their own strengths and weaknesses, and identified a team goal under each dimension. In addition, team members rated their own performance on each of the teamwork dimensions using the same behavioral anchors. This data was analyzed and compared to the raters' assessment of the team. The results were debriefed to the team prior to the next scenario exercise. The team's perceptions of stress level (in terms of workload and level of ambiguity) of the scenario was assessed. (This data can be used to build new scenarios or modify existing scenarios.)
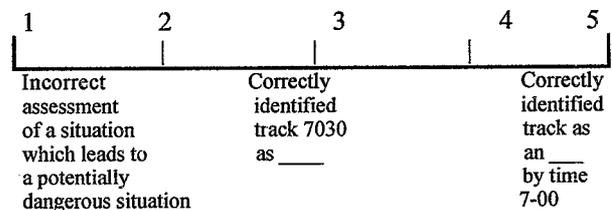
1	2	3	4	5
Incorrect assessment of a situation which leads to a potentially dangerous situation		Correctly identified track 7030 as ___		Correctly identified track as an ___ by time 7-00

Figure 1. Behavioral anchors for situation assessment.

During the debrief, feedback was organized around the trigger events. First, the debriefer defined the teamwork process to be addressed. Sample dialogue was provided to assist the debriefer (refer to Table 2).

Table 2

Sample Dialogue for Situation Assessment

Say: (Sample Dialogue) We will start off by discussing your team's ability to effectively assess each of the critical situations or trigger events in the scenario. What I mean by effective situation assessment is communication that promotes a team awareness of the surrounding environment, both internal and external to the team. This involves timely and accurate reporting of deviations and/or potential problems. As I describe each of the critical events in this scenario, I would like to discuss with you how your team assessed tracks of interest.

By providing sample dialogue and cues, the workload requirements for the preparers of the debrief were reduced significantly. The debriefer cited specific behaviors that were observed and recorded during the event. The team received both positive feedback and constructive criticism. The debriefer described how the targeted behaviors under each teamwork area contributed to the team's performance and the outcomes. Suggested questions or probes were provided to the debriefer in order to solicit the team's self-assessment of their own performance on each of the team process measures (refer to Table 3).

Rater's Training

It is important to train raters in the debriefing methodology, and in how to prepare an effective debrief. Training for those who participated in BFTT's DT-IIA TEV test was designed by the NAWCTSD TADMUS program. For the TEV, raters became familiar with the material by going through the scenario briefing materials, and the On-line ATOM and Debriefing Guide for each scenario. Video tapes of effective and ineffective teams in laboratory settings were used to demonstrate targeted teamwork behaviors. These tapes were also used during practice sessions which provided an opportunity for the raters and designated debriefer to practice using the On-line ATOM worksheets, and to develop mock briefs. Following each trigger event the video tape was paused, and the raters discussed their observations, and how they would classify them using the behavioral anchors. Next, the raters practiced evaluating teams without pausing after each trigger event. With aid of a facilitator, the TEV team rehearsed the process of preparing a debrief. Each rater independently assigned ratings using a 5-point behavioral anchors. Through the process of dialogue and consensus, the ratings were discussed and discrepancies resolved. By the end of the training, the independent ratings never differed by more than one point. The TEV team practiced the debriefing methodology, and a designated debriefer rehearsed providing mock debriefs using the Debriefing Guide format.

Table 3

Sample from Debriefing Guide Worksheet

Say: Event One lasted from the start of the scenario until time 13-00. At the start of the scenario you had track 7030 performing touch and goes at Khark Island.

Ask: What was this track identified as?
Point Out: According to my note,

___ noted aircraft performing touch and goes at ___	___ correctly identified aircraft as ___ at time ___
or	
___ misidentified the aircraft as ___	

Ask: Why was the aircraft misidentified?
Say: Track 7030 could have identified as an ___ by ESM detection of an ___ .

Ask: How was track 7030 evaluated?
Point Out: According to my notes,

___ evaluated 7030 as: ___ nonthreatening at time ___ or ___ threatening at time ___

Ask: What facts went into your decision to evaluate this track as threatening/nonthreatening?
Point Out: The track had a hostile emitter however, it was not ___ and was flying at a low speed.

Ask: What available sources of information did your team use to assess key tracks?
Point Out: According to my notes,

+ Sought pertinent information from available sources

Ask: What additional sources of information were available that you could have used to assess the key tracks?

GUIDELINES

The following guidelines were derived from lessons learned during the BFTT TEV developmental test. Many of these guidelines are supported by other research efforts. The section below provides guidelines for designing: (a) trigger event-based training scenarios, (b) measures of team performance, (c) a performance feedback system, and (d) training for assessing and delivering feedback to facilitate teamwork skills. Future implications will also be discussed.

Scenario Design & Development

The trigger event-based structure used in the BFTT DT-IIA TEV test proved to be effective in eliciting teamwork behaviors. Based on the TADMUS research and the BFTT DT-IIA TEV test results, the following guidelines are offered in the design and development of training simulator scenarios.

- **Stimuli used to develop specific triggers must be requirements/objective based.**

  The identification of training objectives in terms of individual task skills and teamwork skills assists in the development of trigger events designed to elicit specific responses by the team and by individual team members. Identification of the training requirements helps in the design and development of an on-line performance measurement tool.

- **Calibrate the level of difficulty and complexity of the scenario.**

  The scenarios used in DT-IIA were calibrated in terms of the level of stress. Using the same trigger events, two versions of the same scenario were developed. By increasing the workload and the amount of ambiguity in the scenario, high stress conditions were simulated. As a result, individual task skill behaviors and teamwork behaviors could be evaluated under moderate and high stress conditions. Performance differences and the impact of stress on task and team skills could be evaluated.

- **Develop a time line, specifying when a trigger event begins and ends, and what stimuli initiate the actions to be taken.**

  The use of a time line, and specific scripted stimuli (i.e., communications to be delivered at a specific time) cue the raters to the initiation of a key event. This facilitates the on-line capturing and recording of individual and team performance measures.

- **Prior to introducing the first trigger event, provide a minimum of 15 minutes 'lead-in' time before introducing the first tactically significant trigger event.**

  Given the time constraints of the ship's schedule, the scenarios used in DT-IIA were thirty minutes in duration. The scenarios contained three events which averaged 10 minutes in length. Introduction of the first trigger event shortly after the beginning of the training exercise did not allow the team to develop sufficient scenario situation awareness prior to handling the first tactically significant trigger event. During normal ship routines, a CIC team has an opportunity to acquire an awareness of the situation. By providing 'lead in' time, the team can acquire scenario situation awareness. Feedback from the team during the BFTT TEV strongly supports the need to provide sufficient 'lead-in' time for the team to become familiar with the synthetic theater of war before introducing the first tactically significant event.

- **Incorporate multiple trigger events in a training scenario.**

  The use of multiple events within a scenario provides the team and individual team members with more than one opportunity to practice critical task skills and teamwork skills. By incorporating multiple events within a single scenario, raters are able to evaluate the team's skills under a variety of conditions and situations. Performance during one event should not be dependent upon another event within the scenario.

- **Following the last trigger event, allow an additional 15 to 20 minutes 'wrap-up' time.**

  Given the constraints of the ship's schedule, the scenarios used during DT-IIA were limited to 30 minutes. Based on feedback from team members, the team was not given enough time to prosecute the actions taken as a result of the last trigger event. Team members suggested that additional time should be added to the end of a scenario to allow the team sufficient time to execute actions initiated by last event in the scenario.

Team Performance Measurement

The structure and format of the On-line ATOM was well received by combat fleet trainers. It provided a detailed time line for each of the key events with a list of triggers associated with the event. As demonstrated in the conduct of the BFTT TEV, the design of the on-line performance measurement tool facilitated the capturing of specific individual task skills and teamwork measures. The On-line ATOM is sensitive to capturing team process measures that are inherent in highly productive teams. The following guidelines are for the design of an on-line performance measurement tool:

- **To develop behavioral anchors for each trigger event, clearly define critical task skill behaviors and teamwork behaviors that constitute both effective and ineffective performance.**

  Based on the learning objectives and the task requirements outlined by the training scenario structure, specific task skills and teamwork behaviors must be defined. These data are be used to develop (a) a rating system (i.e., refer back to Figure 1) for each of the teamwork dimensions, (b) dialogue for the debriefing guide and (c) cues for the raters.

- **Provide an event time line to record observations of team processes (refer to Figure 2. In Figure 2, SA = Situation Assessment; Comm = Communications; CB = Compensatory Behaviors; and TL = Team Leadership).**

The use of the scenario time line with embedded triggers by specific track numbers facilitates on-line data capture of teamwork skills by shipboard trainers. The structure of the On-line ATOM allowed the raters to rapidly acquire scenario situation awareness, and to anticipate and to look for specific task skills and teamwork behaviors during the conduct of the TEV test.

The On-line ATOM and Debriefing Guide can be easily adapted to new training scenarios. Reactions from the fleet trainers in DT-IIA were positive regarding the practical use and adaptation of the ATOM and Debriefing Guide, and the tool's sensitivity to capture and assess effective/ineffective teamwork skills.

Begin Event One	SA	Comm	CB	TL
00-00 Detected TACAIR "A" at KHARK 0044/38				
00-30 Lost contact on TACAIR 7030				

Figure 2. Sample of an ATOM worksheet.

Debriefing Guide

Fleet personnel who participated as raters and debriefer during the BFTT DT-IIA TEV found the On-line ATOM and Debriefing Guide to be extremely useful diagnostic tools for measuring teamwork measures. Structuring the debrief around the trigger events facilitated debrief preparation, and helped to focus the team on specific behaviors and actions that should have been taken given the simulated synthetic theater of war. The AAW CIC team that participated in the training exercises indicated that the content and structure of the debrief was very useful in identifying ways to improve their team's performance. The following guidelines for debrief are:

• **Through the process of dialogue and consensus, synthesize and integrate data.**

Using the format of the Debriefing Guide, data collected on-line can be synthesized and integrated within 15 minutes. Using a process of dialogue and consensus, a diagnostic evaluation of a team's performance can be prepared. Those who participated in the development of debriefs during the TEV found this process to be very effective in developing a process-oriented debrief.

• **Facilitate learning by creating an environment by which each member of the team becomes an active participant**

The methodology used during DT-IIA created an environment where the team became actively involved in the learning experience. Through active exploration of what had occurred during each of the scenario events (ground truth), the team and the debriefer discussed the specific behaviors observed, and the specific task skill behaviors and teamwork behaviors needed to be an effective CIC team. The team became actively involved in the learning process.

Through the process of dialogue a team can learn "to think together -- not just in the sense of analyzing a shared problem, but in the sense of surfacing fundamental assumptions and gaining insight into why they arise (Isaacs, 1995, p. 26)."

• **Interactive dialogue between members of the team and the raters should focus on those targeted task skills and teamwork behaviors that a team demonstrated or should have demonstrated during each of the trigger events.**

The Debriefing Guide organizes the feedback given to a team around the trigger events. This organizational structure focuses the team's attention on ground truth and helps a team to focus on critical task skills and teamwork skills. The guide helps a team to identify (a) their strengths, (b) areas of opportunity for improvements and (c) team goals to improve performance. Soliciting team input as to why certain actions were taken or omitted helps to establish an interactive dialogue between team members and the debriefer.

• **To reduce the workload of shipboard trainers in preparing a debrief, utilize a standardized on-line performance measurement tool and debriefing guide.**

As demonstrated during the conduct of the TEV test, the workload of the shipboard trainers to

collect performance measures real-time, and to prepare a debrief was reduced. Standard performance metrics (task skills & teamwork skills) for each scenario were used throughout the training exercise and a diagnostic debrief was developed within 15 minutes. Using a standardized performance measurement tool and debriefing format established feedback consistency across training exercises.

Rater's Training

The Rater's Training used in the DT-IIA TEV prepared the raters to evaluate teamwork performance, and to generate and deliver a process-oriented debrief. Reactions from the participants were very positive. The training yielded reasonably consistent ratings across raters from various Navy fleet training groups as well as civilians with combat systems experience. The following guidelines for designing rater's training are offered based on the DT-IIA TEV test:

- **Familiarize raters with the critical trigger events within each scenario.**

- **Provide information regarding scenario trigger events, and target task skills and teamwork behaviors.**

- **Observe teams demonstrating effective and ineffective teamwork.**

- **Practice identifying target behaviors and assigning ratings.**

- **Provide feedback on the accuracy of the rater's assessments.**

Future Implications

Based on this effort, the use of prescriptive, diagnostic tools and a standardized debriefing format was found to be extremely beneficial in highlighting key task skills and teamwork behaviors. These tools generate behaviorally based evaluations and feedback products. The methodology used during DT-IIA establishes and maintains training consistency through standardized evaluation and feedback.

This effort lays the research groundwork for developing and designing support systems for the Combat Systems Training Team (CSTT). Further research is needed to identify methods and technologies that will improve the ability of the shipboard instructor to perform crucial instructional tasks (e.g., guidance for diagnosing performance, implementing instructional strategies, debriefing performance), and to provide instructional aids (i.e., a library of objective based scenario events, automated performance evaluation to and data debriefing products) to assist the shipboard CSTT in (a) designing simulator training scenarios, (b) assessing team and task performance and (c) developing and delivering an effective debrief.

References

Battle force tactical training (BFTT) system developmental test II-A training plan. (1994, July 29). Arlington, VA: Department of the Navy, Naval Sea Systems Command.

Cannon-Bowers, J.A., & Salas, E. (in press). Team performance measurement in training: A conceptual approach to dynamic assessment. To appear in M.T. Brannick, E. Salas, & C. Prince (Eds.), *Assessment and measurement of team performance: Theory, research & applications.* Hillsdale, NJ: LEA.

Cannon-Bowers, J.A., Salas, E., & Grossman, J.D. (1991, June). *Improving tactical decision-making under stress: Research directions and applied implications.* Paper presented at the International Applied Military Psychology Symposium, Stockholm, Sweden.

Cannon-Bowers, J.A., Tannenbaum, S.I., Salas, E., & Volpe, C.E. (1995). Defining team competencies and establishing team training requirements. In R. Guzzo and E. Salas (Eds.), *Team effectiveness and decision making in organizations* (pp. 333-380). San Francisco, CA: Jossey Bass.

Dwyer, D.J., Hall, J.K., Volpe, C.E., Cannon-Bowers, J.A., & Salas, E. (1992). *A performance assessment task for examining tactical decision making under stress* (Special Report 92-002). Orlando, FL: Naval Training Systems Center.

Dwyer, D.J. (1992). An index for measuring naval team performance. *Proceeding of the Human Factors Society 36th Annual Meeting, 2,* 1356-1360.

Entin, E.E., Serfaty, D., & Decker, J.C. (1994, Oct). *Team adaptation and coordination training.* Burlington, MA: AlphaTech, Inc.

Hall, J.K., Dwyer, D.J., Cannon-Bowers, J.A., Salas, E., & Volpe, C.E. (1993). Toward assessing team tactical decision-making under stress: The development of a methodology for structuring team training scenarios. *Proceedings of the 15th Annual Interservice/Industry Training Conference* (pp. 87-98). Washington, D.C.: National Security Industrial Association.

Isaacs, W. N. (1995, May). Taking flight: Dialogue, collective thinking, and organizational learning, *Organizational Dynamics,* 24-39.

Johnston, J.H., Cannon-Bowers, J.A., & Smith-Jentsch, K.A. (1995). *Event based Performance measurement system for shipboard command teams.* Paper presented at the 1995 Proceedings of the First International Symposium on Command and Control Research and Technology.

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

Prince, C., Oser, R., Salas, E., & Woodruff, W. (1993). Increasing hits and reducing misses in CRM/LOS scenarios: Guidelines for simulator scenario development. *The International Journal of Aviation Psychology, 3,* 69-82.

Salas, E., Cannon-Bowers, J.A., & Johnston, J.H. (in press). How do you turn a team of experts into an expert team? Emerging training strategies. To appear in C. E. Zsambok & G. A. Klein (Eds.), *Naturalistic decision making.* Hillsdale: NJ: LEA.