# Measures for Evaluation of Team Processes and Performance in Experiments and Exercises

Elliot E. Entin and Eileen B. Entin
Aptima, Inc.
12 Gill Street, Suite 1400
Woburn, MA 01801
781.935.3966, ext. 238
Entin@Aptima.com

## Abstract

A critical need exists for a solid understanding of the factors that influence team decision making and performance in order to identify interventions that can affect the decision making process and improve performance (Klinger et al., 1993; Salas, Bowers, and Cannon-Bowers, 1995; Hall and Regian, 1996). In this paper we describe a variety of team-based measures we have used to assess adaptive teams in simulation environments. The measures reflect quantitatively-based assessments obtained from observers and participants in simulator-based scenarios, and complement those obtained from the simulator. They have been applied in training and evaluation exercises and in research programs across a variety of domains, and have provided valuable information in all their applications. In the Adaptive Architectures for Command and Control (A2C2) program, they have allowed us to characterize and compare team performance between traditional and adaptive command and control architectures, and to compare actual team performance to model-based predictions.

## Introduction

A critical need exists for a solid understanding of the factors that influence team decision-making and performance in order to identify and manipulate interventions to affect the decision making process and improve performance (Klinger et al., 1993; Salas, Bowers, and Cannon-Bowers, 1995; Hall and Regian, 1996). In this paper we describe a set of measurement approaches used by various research programs, including the Tactical Decision Making Under Stress **(**TADMUS) and the Adaptive Architectures for Command and Control (A2C2) programs, to characterize and analyze team performance and processes in simulator-based exercises. We focus on measures that can be obtained from observers and participants, which are independent of the simulation environment in which the exercise or training is conducted. Measures obtained from the simulation environment will complement or provide an alternative view of the observer- and participant- based measures. These measurement approaches have been applied in training and evaluation exercises, and in research programs, and have provided valuable information in all their applications. Similarly, the measures have been used successfully across a variety of application domains.

Observer-based measures are obtained from knowledgeable, trained observers who observe team performance and processes during a scenario run. The ratings they provide can pertain to individual members of the team, to subcomponents of the team, or to the team as a whole. Participant-based measures can focus on self, other team members (individually or collectively), and team as a whole. In most cases, participant-based measures are typically obtained at the end of a scenario, but in some cases it is possible to halt the simulation at specified times to obtain interim assessments.

Table 1 provides a high-level view of observer- and participant-based team performance, process, and climate measures that have been used in team research projects. The four columns in the table show, respectively, the name of the measure, the level at which it is captured (individual or team), the source of the information (observer or participant), and an explanation of the measure. These measures can be obtained in virtually all simulator-based exercises. The particular measures that are collected in any given situation will depend upon the factors incorporated into the scenario and the goals of the experimental or training context in which the measures are collected.

In the next two sections we provide more detailed explanations of the measures in Table 1, the instruments that are used to collect the data, and the method by which the data are collected. We look first at observer-based measures and then at participant-based measures. For observer-based measures, we note if training is required. In the case of participant-based measures, we note any explanatory materials that must be provided before participants complete the data collection instruments.

## Observer-Based Measures

### *Team Performance Outcome*

The team performance outcome measure is comprised of a set of behaviorally-anchored scales that capture the quality of the team's performance on a mission. To apply the measure, which is based on the Anti-Air Warfare Team Performance Index (Johnston, Smith-Jentsch, and Cannon-Bowers, 1997), it must be tailored to a particular scenario and mission. A subject-matter expert decomposes the mission into its component team tasks and, for each task, specifies as the behavioral anchors the key behaviors that would indicate superior performance, adequate performance, and poor performance. Items in the questionnaire address the team's performance on each of these tasks, as well as on overall performance. An example of a performance outcome item for a scenario involving a joint task force engaged in establishing a forced insertion is shown in Figure 1. Entin, Serfaty, and Deckert (1994), using a performance outcome measure composed of 12 items, showed that the team performance outcome measure possessed high reliability and high construct validity.

The key behaviors associated with each task provide the basis for a behaviorally anchored 7-point scale for rating performance on the task dimensions. When the measure is used in a simulated mission, one or more trained observers watch the team as it performs its mission, and take notes on its performance. At the end of the mission, using the behaviorally anchored scales, the observers rate the team on the performance of each of the tasks and on its overall performance on the mission.

Table 1. Overview of Observer-based and Participant-based Measures

| **Measure** | **Level (focus) of observation** | **Source of data** | **Description** |
|---|---|---|---|
| Performance Outcome | team | Observer | Behaviorally anchored ratings of quality of aspects of and overall team performance |
| Teamwork | team | Observer | Behaviorally anchored ratings of quality of six dimensions of teamwork processes: communication, monitoring, feedback, back-up, coordination, and team orientation. |
| Team Processes and Dynamics | Individual, team | Participant | Enumeration of unobservable individual and team factors underlying team processes derived from scenario-based structured interviews |
| Verbal Communication | Individual | Observer | Records of type, sender and recipient, type and time of communications (see Table 2 for description of measures) |
| Workload | Individual, team | Participant | Assessment of individual workload for self and others; global assessment of team workload |
| Mutual mental model congruence | Individual, team | Participant | Assessment of the congruence of models team members hold of one another |
| Organizational awareness | Individual | Participant | Assessment of the accuracy or congruence of team members' situational and mutual mental models |
| Scenario and tools evaluation | Individual, team | Participant | Ratings of aspects of scenario including level of difficulty, complexity, uncertainty, ambiguity for self, others, and/or team as a whole |
| Attitude/climate evaluation | Individual, team | Participant | Ratings of attitudes, feelings, and opinions pertaining to selected issues or topics |

Clear SAMs. Pertains to destroying the SAM installation around the airfield and the port.

| Very Poor | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Superior |
|---|---|---|---|---|---|---|---|---|

7 Team brought appropriate assets (e.g., 1 CAS + SOF or 1 CAS + SAT) to bear to destroy SAM installations around the airfield and the port in an efficient and timely fashion. The attacks were carried out in sufficient time prior to attacking the airfield and port so there were no delays.

1 Team failed to destroy SAM installations around the airfield and the port, used inappropriate assets, and/or were inefficient or not timely. The attacks on the SAMs were late such that attacks on the airfield and port were held up or Blue forces took casualties.

Comments:______________________________________________________

Figure 1. Example Performance Outcome Item with Behaviorally Anchored Scales

With behaviorally anchored scales observers have fixed criteria against which to rate a team's performance. In contrast, a relative low to high rating scale is sufficient for comparing one set of teams against another set, or for comparing team performance under different conditions, but does not allow for comparison to an absolute standard. The performance assessment measure can be used to assess team performance on an absolute scale, a capability that is particularly important in situations such as a team training exercise, where there is no comparison group or condition against which a team's performance is being measured.

Analysts can use this measure to look at team performance on specific tasks, subsets of tasks, as well as on the overall mission. There are two alternative metrics to characterize overall performance: the arithmetic mean of the tasks and the overall performance rating supplied by the observer. Typically these two measures of overall performance are highly correlated. Subscales can be used to assess consistency in performance across the mission tasks.

The performance outcome measure was originally developed under and used extensively in the Navy-sponsored TADMUS program (Cannon-Bowers and Salas, 1998). The measure has been extended, refined, and extensively applied in the A2C2 program. In this research program a series of experiments was designed and conducted to evaluate the effectiveness of alternative team architectures and varying mission conditions (Entin, Serfaty, and Kerrigan, 1998; Entin, 1999, 2000; Entin, Entin, and Serfaty, 2000). In all experiments the performance outcome measure was positively correlated with performance metrics obtained from the simulator (Entin, Serfaty, and Kerrigan, 1998; Entin, 1999). Scores on the individual items are in most cases moderately correlated and each is predictive of the overall scale score. The individual item scores can be used as a diagnostic and feedback tool for specific task behaviors, and they can be related to the various team process measures that are used.

*Teamwork*

The teamwork measure has its roots in the anti-air teamwork observation measure (ATOM; Smith, 1994). The original measure was comprised of 15 items that measured six dimension of teamwork: communication, monitoring, feedback, back-up, coordination, and team orientation. Based on extensive factor and construct analyses (Entin and Serfaty, 1995), the instrument has been distilled to a smaller set of items that capture the quality of a team's teamwork processes on the six dimensions. The teamwork measure is independent of the task domain and mission objectives associated with any particular scenario. Thus, the same instrument can be used in a variety of applications, and there is no requirement for input from a subject-matter expert in developing it (as there is with the team performance measure). The six dimensions of teamwork can be examined individually (for example, to look specifically at team back-up behavior), in subsets, or as a whole. The arithmetic average of the six ratings provides an overall measure of team process.

The team process measure is captured by observers, who rate each of the six dimensions on a behaviorally anchored 7-point scale. For each dimension, at one end of the scale are examples of behaviors indicating poor team process (for example, poor monitoring behavior), and at the other end are behaviors indicative of good team process on that dimension. An example item from the teamwork measure is shown in Figure 2. During the scenario run, observers take notes on team processes, and at the end of the scenario they complete the behavioral ratings, based on their observations across the entire scenario.

**Monitoring Behavior**

To what extent did team members alert each other to impending decisions and actions?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

7 Team members always alerted each other to impending decisions and actions; supporting information was actively solicited from other team members.

1 Team members did not keep each other informed of impending decisions and actions; compromises to mission safety or mission effectiveness arose when a team member waited for the other to volunteer significant information.

Comments:____________________________________________

Figure 2. Example Teamwork Item Illustrating Behaviorally Anchored Scale

As is the case with the team performance measure, the teamwork measure also uses a fixed, behaviorally-anchored rating scale, so that the teamwork processes exhibited by a team in a given simulation can be evaluated against a fixed standard, rather than compared relative to another group or another condition. Again this capability is particularly important in practice or training situations in which the concern is focused on absolute, rather than relative, quality of performance.

The teamwork measure provides a way of understanding team performance. It was used in the TADMUS program to assess the effectiveness of a Team Adaptation and Coordination Training (TACT) program (Entin and Serfaty, 1999). It has been used in the A2C2 program to analyze and understand how alternative team architectures facilitate or inhibit effective team processes (Entin, Serfaty, and Kerrigan,1998; Entin, 1999, Entin, Entin, and Serfaty, 2000). As with the performance outcome measure, the individual scales comprising the teamwork measure can be used to provide teams with feedback as to which teamwork dimensions require improvement.

In order to insure that the ratings are reliable, observers who have not used behaviorally anchored scales should receive training before they use the team performance outcome and teamwork measures for the first time. Because the assessment of team performance outcome and teamwork occur at the same time (typically at the end of the scenario), because both measures rely upon behaviorally-anchored scales, and because they are carried out by the same observer(s), training for both instruments can be conducted jointly.

*Team Processes and Dynamics*

The teamwork measure captures team processes based on behaviors that can be observed during the execution of a scenario. To understand and describe more completely the factors that entered into individual and team processes, decisions, and actions, we conduct scenario-based structured interviews with individual members of a team or with the team as a whole. In these interviews, we focus on a particular event or time in the execution of a scenario and pose a set of questions to understand matters such as what factors the individuals were considering at that time, what pieces of information they were weighing, what information they were seeking, and why they made a particular decision or took a particular action. The particular questions that are posed are determined both by the nature of the scenario and the purpose of the investigation. For example, if the goal is to understand what pieces of information team members considered in arriving at a particular decision, the questions might be: What led up to the situation? What recourses are involved? What resources might the other team members need? How do you (the decision maker) perceive the situation? What are the expected outcomes of the decisions? Were the expected outcomes achieved?

Entin (1996) used structured interviews to solicit the views of general officers on current and future joint command and control, and adaptive architectures. This measurement approach was especially useful for eliciting the officers' views on the drivers of adaptation.

*Team Communications*

When communications among the team members are captured at a detailed semantic level, it is very difficult and time consuming to develop meaningful, quantitatively-based measures to describe the nature of the communications. On the other extreme, simple frequency counts of communications, though straightforward, do not provide a meaningful window into team processes. In an Aptima-developed approach, verbal communications among the members of the team are captured by observers at an *intermediate* level of detail that incorporates both semantic and quantitative aspects of the communication stream.

Using the Aptima-developed technique, during the run of a scenario observers listen to the communications in real time, and use a specially designed form to code the source, the recipient, the time (if a hand-held computer is used), and the type of the verbal communications among the team members. Types of communications are divided into three basic categories: transfers, requests, and acknowledgements. Both transfers and requests, in turn, can be classified as requests for information, action, or coordination.

The recording form is implemented in a matrix format. An example of a communication matrix is shown in Figure 3. This instantiation of the matrix is one that was used in the A2C2 (Entin, Serfaty, and Kerrigan, 1998; Entin, 1999) research programs. This is a portion of the observation sheet for one observer. The darkened cells were not being observed by the observer using this particular form of the matrix.

Team ID________ Trial ________ Date__________ Observer______________ Part I Team

| Type & Content | | FLAG TO: | | | | | GREEN TO: | | | | | BLUE TO: | | | | | Type |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | GREEN | BLUE | PURPLE | RED | ORANGE | FLAG | BLUE | PURPLE | RED | ORANGE | FLAG | GREEN | PURPLE | RED | ORANGE | |
| Request | Informatio | | | | | | | | | | | | | | | | Request |
| | Action & Task | | | | | | | | | | | | | | | | |
| | Resource Utilization | | | | | | | | | | | | | | | | |
| | Coordination | | | | | | | | | | | | | | | | |
| Transfers | Information | | | | | | | | | | | | | | | | Transfers |
| | Will Perform Action / Task | | | | | | | | | | | | | | | | |
| | Will Use Resource | | | | | | | | | | | | | | | | |
| | Will Coordinate | | | | | | | | | | | | | | | | |
| Acknowledge-ments | | | | | | | | | | | | | | | | | Ack |

Figure 3. Example of Matrix Used to Capture Team Verbal Communication

The recording form was originally developed as a paper-and-pencil instrument, but has more recently been implemented on a hand-held computer. With the computer-based version time stamped observations can be obtained and an associated post processing routine can be used to derive a number of different communication measures that reflect the quantity, directionality, timing, and type of communications that occur. Coordination measures can also be derived from the data that is collected. In order to compare scenarios that may be of different durations, most measures are calculated per unit time, typically per minute, rather than as raw totals.

Table 2 explains some of the most common communications measures that are derived from the recording matrix. These measures can be captured at the individual level or at the team level. In addition, in hierarchical teams, they can be captured at the leader and subordinate levels, and the results can be used to look at upward, downward, and lateral communication.

Table 2. Measures of Verbal Communication Derived from Communications Matrix

| **Measure** | **Description** |
|---|---|
| **Overall Rate** | |
| Total Communications | Total number of communications per minute |
| **Communication Types** | |
| Information Requests | Number of requests for information per minute |
| Information Transfers | Number of transmissions of information per minute |
| Action Requests | Number of requests for an action per minute |
| Action Transfers | Number of statements of actions (to be) taken per minute |
| Coordination Requests | Number of requests to coordinate an action per minute |
| Coordination Transfers | Number of agreements to coordinate an action per minute |
| Acknowledgements | Number of non-substantive acknowledgements of receipt of communication (e.g., ‘ok’ to acknowledge receipt of information) per minute |
| **Communication Ratios** | |
| Overall anticipation | All communication transfers divided by all communication requests |
| Information anticipation | Information transfers divided by information requests |
| Action anticipation | Action transfers divided by actions requests |

Anticipation ratios, measures based on the ratio of two measures, have often proved more useful than individual rate measures for understanding team communication (Entin, Serfaty, and Deckert, 1994; Entin and Serfaty, 1999; Serfaty, Entin, and Johnston, 1998). The information anticipation ratio, the ratio of information transfers to information requests, has proved particularly useful for understanding team communications. Ratios larger than 1.0 are assumed to indicate that team members are anticipating the information needs and requirements of other team members and pushing them information before they request it (Entin, Serfaty, and Deckert, 1994). Ratios less than 1.0 are assumed to indicate that little anticipation of information needs is occurring and team members must request (pull) the information they require from others.

The communication measures have proved useful throughout the A2C2 program for analyzing the efficiency and effectiveness of alternative team architectures (Entin, Serfaty, and Kerrigan, 1998; Entin, 1999; Entin, Entin, and Serfaty 2000). A version of the communication matrix was applied in an Army evaluation exercise conducted at Fort Sill, OK. This application provided an example demonstrating that the verbal communications matrix could be adapted and team communications recorded in an exercise involving teams of teams. In this situation, communications were sampled in selected intervals of approximately 35 minutes duration. Ten samples were taken over a 5-day period, providing data for comparisons of team communication patterns over time and across activities.

In order to insure that team communications are captured reliably, it is necessary to provide observers with training and experience in using the communications recording matrix. This entails coding short scenario to practice using the categories and becoming familiar with the speed at which events occur in a prototypical scenario. Training to reach an acceptable level of proficiency usually takes 25 to 30 hours.

## Participant-based Measures

*Team Workload*

The Task Load Index (TLX; Hart & Staveland, 1988) provides an assessment of individual team members' workload. Aptima has extended the TLX to capture team as well as individual workload (Entin, Serfaty, and Kerrigan, 1998). We use a three-part questionnaire to assess individual and team workload. In the first part of the workload questionnaire, participants report their own workload in terms of five of the traditional items comprising the TLX[1]. In the second part of the questionnaire each participant provides an estimate of the overall workload experienced by each of the other team members. In the third part of the questionnaire each participant responds to the five TLX items, but this time for the team as a whole (not just for themselves).

Individual workload has been assessed using the TLX in numerous research projects. In a project investigating the impact of information load on information processing, the workload measure was used to verify that the participants experienced a higher level of workload under the high than under the moderate information load condition (Entin, Kerrigan, Serfaty, Klein, and Wolf, 1998). In the A2C2 research program, individual workload has been measured in all experiments (Entin, Serfaty, and Kerrigan, 1998; Entin 1999, 2000).

The team workload measure has been used extensively in various domains. Team workload has been assessed in numerous studies conducted under the TADMUS and A2C2 programs, and has been found to be negatively related to measures of team performance (Entin and Serfaty, 1999, Entin, 1999). An interesting finding that emerged from several A2C2 studies is that, contrary to expectation, individuals consistently rate

[1] We usually omit the sixth item, physical workload, because it is not applicable in most simulation-based situations.

the workload of their team members as being higher than their own (Entin, Serfaty, and Kerrigan, 1998).

*Mutual Mental Model Congruence*

In addition to providing measures of the level of individual and team workload, Aptima uses data from the workload assessment questionnaire in an innovative method for assessing the congruence of the team's perception of workload across members of the team. The measure of congruence provides a window into the accuracy of the team members' mutual mental models of the team functioning. The Adaptive Team Model (Serfaty, Entin, and Johnston, 1998) suggests that the development of shared situational mental models of the task environment, the task itself, and of interacting team members' tasks and abilities are used to generate expectations about how other team members will behave (McIntyre and Salas,1995; Cannon-Bowers, Salas, and Converse, 1993); Orasanu, 1990; Serfaty, Entin and Volpe, 1993). There is an increasing amount of research evidence showing that high-performing teams make use of mutual mental models (Cannon-Bower, Salas, and Converse, 1993; Entin and Serfaty, 1999; McIntypre and Salas, 1995; Orasanu, 1990).

To evaluate the accuracy with which team members can estimate the individual workload of each of their fellow team members, we calculate congruence or deviation measures that reflect the difference between a team member's self-reported workload and the estimates of his or her workload made by each other team member. One such deviation measure is the root mean square. To compute this measure: 1) the self-reported workload for an individual is subtracted from each team member's estimate of the workload for that individual; 2) these difference scores are squared, summed, and averaged for the team, and; 3) the square root of the average is taken. The mutual mental model congruence measure has been shown to have predictive validity in that it covaries with performance outcome and teamwork (Entin, 1999).

*Organizational Awareness*

Aptima has developed an innovative measure of organizational awareness that provides a window into the coherence of the team's situational and internal mental models. We conceptualize team members' knowledge of the roles, tasks, and relationships of other team members as one aspect of their mental model of how the team is structured and how it functions. The organizational awareness measure assesses the congruence of team members' perceptions of team member activities at a given time during a scenario.

Data for the organizational awareness measure is typically gathered at the end of a scenario run. A questionnaire requests each team member to retrospect at the end of a scenario about one or more salient events within the scenario. Each event acts as a common time marker so that all team members are focused at the same time within the scenario. For each event each team member first reports the task he or she was performing when the event occurred and then reports the task he or she believes each of the other team members was performing during the same event. To develop the organizational awareness measure, the task category representing what each team member said he or she was doing is compared to the task category representing what each of the other team members said that team member was doing. The number of category matches is counted and a percentage agreement (congruence score) computed for each team.

In order to use this measure, it is necessary to identify and focus attention on one or more events that are salient for all team members, and ones that are almost certain to occur in every scenario trial. For example, if a scenario portrayed two waves of enemy attacks, a salient event might be the start of the second attack. If, however, the occurrence of the second wave were dependent upon the way in which the scenario unfolded in real time, this would not be a suitable event on which to focus.

Results from a series of studies conducted under the A2C2 program show that the organizational awareness measure is related to other team process measures and to performance as well, confirming that it is a valid approach for assessing the congruence of a team's mental model (Entin and Entin, 2000). To capture the hierarchical nature of a team one can calculate a similar, but more narrowly focused, score based on the team leader's knowledge of what the other team members are doing and, conversely, the team members' knowledge of what the leader is doing. Discrepancies in organizational awareness across hierarchical levels of an organization may be diagnostic of an organization's poor performance.

*Scenario Assessment*

At the end of a scenario run we may administer a questionnaire to the participants that asks for their assessment of certain characteristics of the scenario such as its complexity, ambiguity, uncertainty, and realism. The questionnaire may also ask team members to provide information about aspects of team performance and processes. It can also include items to assess whether participants' expectations about the training or practice trial have been met, the degree to which they were successful in using any new, advanced information technology tools that were available, and the degree to which they found those tools helpful in performing their mission. Post-scenario questionnaires typically contain a combination of rating scales, check lists, and short answer questions.

A post-mission questionnaire used by Entin, Serfaty, and Deckert (1994) in an experiment assessing the effectiveness of a team adaptation and coordination training (TACT) program was shown to have excellent reliability. A discriminant analysis suggested that the measure was successful in discriminating between teams that received the TACT training and those that did not (Entin and Serfaty, 1995).

*Attitude/Climate Survey*

At the end of a scenario trial and/or at the end of an evaluation session, we can administer a questionnaire to the participants that asks them to assess individual and/or team attitudes about aspects of team climate such as the individual or team motivation, cooperation, support, and level of functioning. The post-scenario questionnaire can also incorporate questions about the extent to which teamwork processes and the team environment in the simulation environment seemed to be consistent with what might occur in an operational environment, or the extent to which team members were actively engaged in the scenario. The particular items included in the questionnaire are determined by the purpose of the session that is being evaluated, and may be short answer questions, rating scales, check lists, or items to be ranked on some criterion.

In the evaluation exercise conducted at Fort Sill, OK, a questionnaire administered on the last day of play to teams in the effects cell included items about the adequacy of information availability, and the effectiveness of the organization, the effects cell, and the fire control system being evaluated. Participants' responses to the questionnaire indicated that in order for team members to maintain good situation awareness and plan for future actions, the organizational structure must make it clear where critical information is to be found and must facilitate obtaining that information. The questionnaire also confirmed respondents' belief that good teamwork skills were important to task and mission completion.

## Integrated Applications

The suite of observer and participant based measures that we have described provides a standardized and repeatable methodology for the collection and analysis of data. Use of these measures brings rigor and control to the data gathering process, allows for the application of aspects of scientific methodology in non-laboratory situations, and introduces aspects of reliability and validity into the assessment process. For example, use of the communications matrix form allows trained observers to provide communications data from exercise environments that are consistent and reliable across observers and across sessions as well.

The quantitatively-based measures that we have described can support or augment more qualitatively based descriptive information that exercise observers report. In some cases, the quantitative measures provide support for a specific observation. For example, some of the communications measures described in Table 2 provide quantitative support for observations about team communication and coordination made by exercise observers. In other cases, observations, hypotheses, or explanations of specific situations advanced in AARs by exercise observers or participants can explain or support the findings that emerge from the observer-based or participant-based measures that we have described. For example, in the Global '99 exercise AAR findings indicated that the effects-based operations cell was not adequately addressing the effects issues during game play. Results from the end of exercise survey underscored the AAR findings, and revealed more specific information about why it was not effective. Responses to the questionnaire indicated that effect cell was moving too slowly to be effective, was not establishing necessary procedures to evaluate when an effect had been achieved, and did not establish which agency would do the effects evaluation.

## References

Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In N.J. Castellan, Jr. (Ed.), *Current issues in individual and group decision making* (pp. 221-246). Hillsdale, NJ: Erlbaum.

Cannon-Bowers, J. A. and Salas, E. (1998). *Making decisions under stress: implications for individual and team training*. Washington, D. C. American Psychological Association.

Entin, E. E., Serfaty, D., and Deckert, J. C. (1994). *Team adaptation and coordination training*, TR-648-1. Burlington, MA: ALPHATECH, Inc.

Entin, E. E. (1996). Current and future views on Joint C2 and adaptive architectures. *Proceedings of the 1996 Command and Control Research and Technology Symposium*, Monterey, CA.

Entin, E. E. (1999). Optimized command and control architectures for improved process and performance. *Proceedings of the 1999 Command and Control Research and Technology Symposium*, Newport, RI.

Entin, E. E. (2000). Communication and coordination across low and high fidelity simulations. *Proceedings of the 2000 Command and Control Research and Technology Symposium*, Monterey, CA.

Entin, E. B. & Entin, E. E. (2000). Assessing Team Situation Awareness in Simulated Military Missions. *Proceedings of the Human Factors and Ergonomics Society 44th Annual Meeting*, San Diego, CA.

Entin, E. B., Entin, E. E., and Serfaty, D. (2000). *Organizational structure and adaptation in the joint command and control domain*., TR-915, Burlington, MA: ALPHATECH.

Entin, E. E. & Hess, S. (2000). *The effects of leader role and task load on team performance and process*. A Phase II Report. Woburn, MA; Aptima, Inc.

Entin, E. E., and Serfaty, D. (1995). Team adaptation and coordination training: emerging issues in distributed training. Tr-696, Burlington, MA: ALPHATECH.

Entin, E.E. and Serfaty, D. (1999). Adaptive team coordination. *Human Factors, 41*,2, pp.312-325.

Entin, E.E., Serfaty, D., & Deckert, J.C. (1994) Team adaptation and coordination training (TR-648-1). Burlington, MA: ALPHATECH, Inc.

Entin, E. E., Serfaty, D., and Kerrigan, C. (1998). Choice and performance under three command and control architectures. *Proceedings of the 1998 Command and Control Research and Technology Symposium*, Monterey, CA.

Hall, E. and Regian, W. (1996). Cognitive engineering for team tasks. AFOSR Laboratory Research Initiation Request.

Hart, S. G. & Staveland, L. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Mishkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam: Elsevier.

Johnston, J.H., Smith-Jentsch, K. A., and Cannon-Bowers, J.A. (1997). Performance measurement tools for enhancing team decision making. In M.T. Brannick, E. Salas, and C. Prince (Eds.), *Team performance assessment and measurement: Theory, research, and applications* (pp. 311-330). Hillsdale, NJ: Erlbaum.

Klinger, D. W., Andriole S. J., Militello L. G., Adelman, L., Klein, G., and Gomes, M. E. (1993). *Designing for performance: A cognitive systems engineering approach to modifying an AWACS human computer interface*. Armstrong Laboratory, AL/CF-TR-1993-0093, Wright-Patterson AFB, OH.

McIntyre, R. M., and Salas, E. (1995). Team performance in complex environments: What we have learned so far. In R. Guzzo and E. Salas (Eds.), *Team effectiveness and decision making in organizations* (pp. 9-45). San Francisco: Jossey-Bass.

Orasanu, J. M. (1990). *Shared mental models and crew decision making* (CSL Report 46). Princeton, NJ: Princeton University, Cognitive Science Laboratory.

Salas, E., Bowers, C.A., and Cannon-Bowers, J.A. (1995). Military team research: 10 years of progress. *Military Psychology*, 7, 5575.

Serfaty, D., Entin, E.E., and Johnston, J. (1998). Team Adaptation and Coordination Training. In *Decision Making Under Stress: Implications for Training and Simulation,* Eds. J. A. Cannon-Bowers and E. Salas, Washington D.C.: APA Press.

Serfaty, D., Entin, E.E., and Volpe C. (1993). Adaptation to stress in team decision-making and coordination. *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting.* Santa Monica, CA: Human Factors and Ergonomics Society.

Smith, K. (1994). *Online anti-air warfare teamwork observation measure*: (unpublished paper). Orlando, FL: Naval Air Warfare Center Training System Division.