MOBILE AID FOR TRAINING AND EVALUATION (MATE):
A HAND-HELD, CONFIGURABLE SET OF TEAM PERFORMANCE MEASUREMENT TOOLS

Denise M. Lyons and Robert C. Allen
Naval Air Warfare Center Training Systems Division
Orlando, FL

Abstract

Researchers at the Naval Air Warfare Center Training Systems Division (NAWCTSD) have developed multiple instructor aides for performance measurement hosted on hand-held computers. The initial prototype, Shipboard Mobile Aide for Training and Evaluation (ShipMATE), provides instructors with an automated tool for presenting pre-brief information, collecting data and conducting a debrief utilizing the Team Dimensional Training strategy. This tool was originally presented in 1997. Since then, additional training and evaluation software applications have been developed for Air, Surface, Ground, and Sub-Surface domains in both classroom and operational settings. We refer to the set of available performance measurement tools, which can be loaded onto a hand-held device in any combination, as MATE. The applications on MATE have been tailored to accommodate various performance measurement methodologies, including outcome and process measures, aimed at capturing individual and team performance. Examples of screen designs will illustrate how each software tool is a unique combination of tools such as embedded checklists, a scrolling scenario script window, data tagging buttons, organizational tabs, voice recording, digital handwriting, instructor cueing, links to embedded systems and networking of multiple hand-held devices.

This paper will review the multiple tools that have been developed and present suggested utilization of the optional screen functions. The value added by each function will be discussed as well as the viability of this technology in various operational settings including sea trials aboard AEGIS ships and field operations during Army exercises. Finally, guidelines for developing hand-held instructor aides for future training systems will be summarized.

Biographical Sketch:

Dr. Denise M. Lyons is a Senior Research Engineer in the Science and Technology Division of the Naval Air Warfare Center Training Systems Division. Her current focus includes modeling and simulation, and training technology for programs in Advanced Embedded Training Systems (AETS) and Improved Behavioral Representation of Computer Generated Forces (CGF). She received her Ph.D. in Optical Sciences from the University of Arizona, and has over 13 years of experience in research and development for military systems.

Robert C. Allen is a research assistant in the Science and Technology Division of the Naval Air Warfare Center Training Systems Division. He earned his Masters Degree at New Mexico State University and is currently a Doctoral Candidate in the University of Central Florida's Psychology Department, Human Factors tract. His current focus includes Usability and HCI-related issues and Virtual Environment (VE) research. He has five years experience in VE research and one year experience in Air Traffic Control research.

MOBILE AID FOR TRAINING AND EVALUATION (MATE): A HAND-HELD, CONFIGURABLE SET OF TEAM PERFORMANCE MEASUREMENT TOOLS

Denise M. Lyons and Robert C. Allen
Naval Air Warfare Center Training Systems Division
Orlando, FL

INTRODUCTION

Researchers at the Naval Air Warfare Center Training Systems Division (NAWCTSD) have developed multiple instructor aides for performance measurement hosted on hand-held computers. The initial prototype, Shipboard Mobile Aide for Training and Evaluation (ShipMATE), provided instructors an automated tool for presenting pre-brief information, collecting data and conducting a debrief utilizing the Team Dimensional Training strategy. ShipMATE was originally presented in 1997 (Pruitt, Burns, Wetteland, & Dumestre, 1997). Since then, additional training and evaluation software applications have been developed for Air, Surface, Ground, and Sub-Surface domains in both classroom and operational settings. We refer to the set of available performance measurement tools, which can be loaded onto a hand-held device in any combination, as MATE. The applications on MATE have been tailored to accommodate various performance measurement methodologies, including outcome and process measures, aimed at capturing individual and team performance. Examples of screen designs will illustrate how each software tool is a unique combination of functionality's, such as embedded checklists, scrolling scenario script windows, data tagging buttons, organizational tabs, voice recording, digital handwriting, instructor cueing, links to embedded systems and networking of multiple hand-held devices. Finally, the viability of this technology in an operational setting will be discussed as demonstrated by several diverse field trials.

BACKGROUND

Due to an emphasis on affordability and reduced manning, the military will increasingly rely upon the use of embedded and onboard training systems for maintaining operational readiness. Critical to the success of this strategy is the ability of operators and tactical decision-makers to participate in meaningful and realistic training exercises, coupled with the ability to specify and implement measurement and diagnostic tools for tracking the complex team interactions that underlie team performance. Creating these tools is extremely labor intensive for the developers and instructors who must implement them. Therefore, it becomes imperative that performance measurement solutions tap into emerging technologies that can automate many tasks and alleviate workload.

An example of an emerging technology through which automated performance measurement can be implemented is the Navy's shipboard training environment to enable combat systems team training, called the Battle Force Tactical Training System (BFTT). One goal of the BFTT system is to provide ships' personnel with the ability to self-train and, by doing so, help maintain the continuum of training. Key to the success of this approach is the identification of specific, quantifiable and measurable training objectives. To meet these objectives, Commander in Chief, U.S. Pacific Fleet, has approved the implementation of a fleet training concept called Objective Based Training (OBT).

The OBT process is a systematic approach to training that defines the tasks a ship, team or individual must perform, the manner in which these tasks must be performed, and the standard to which they must be performed. OBT currently is manually implemented within the Afloat Training Organizations. That is, binders of paper based training objectives for each ship class are maintained. These binders contain terminal objectives, enabling objectives and performance measures of various warfare areas, such as Combat Systems, Engineering, Damage Control, Seamanship, Navigation, and Aviation. However, the use of these paper-based tools both during training and post-training analysis is extremely labor intensive. Therefore, the requirement for an automated tool called the Afloat Training Exercise and Management System (ATEAMS) has been developed. Included in the ATEAMS specifications is the need for a hand-held automated performance assessment tool, such as ShipMATE, that can be used to provide meaningful performance feedback to combat systems personnel.

MATE APPLICATIONS

MATE is a computer application containing a suite of tools that can be customized (grouped) to fit a variety of training environments. MATE is loaded onto a Hand-Held Device (HHD) and can be used by trainers to capture observations about team and individual performance in real time during a training exercise. The HHD has infrared and radio frequency communications capability that allows the trainer to

be un-tethered from the training/simulation system (Pruitt, Burns, Wetteland, & Dumestre, 1997).

Some of the major advantages of the MATE training device include:

1. Capturing an instructor's observations about team and individual performance in real-time during a training exercise.

2. Automated time stamping associated with targeted behaviors. The behaviors to be observed/recorded can be pre-specified (scripted and cued) or recorded on the fly.

3. Scripted events linked to performance measures increase the consistency of observations made, and feedback provided, by both expert and novice instructors.

4. Quick data reduction to aid in diagnosis of performance, permitting timely delivery of feedback.

5. Meaningfully organized performance observations, tied with instructor notes and scripted events, facilitates situational awareness during debriefing and improves the effectiveness of feedback given.

6. Digital performance data can be easily archived for trend analysis. Such analysis can be used to target specific deficiencies during the next training opportunity.

The original purpose of ShipMATE was to support Naval instructors when conducting shipboard or land-based training. ShipMATE was designed to aid these trainers in the task of preparing for, conducting, and debriefing objective-oriented, scenario-based training. However, the tools of MATE can be tailored (grouped) to meet the needs of diverse training environments. As a result of this flexibility a variety of applications have been created, each with its own set of objectives and tools used to meet those objectives. The following is a brief description of some of these applications. Tools unique to each application are described in detail.

ATPI
The Air Warfare Team Performance Index (ATPI) was a tool developed under the Tactical Decision Making Under Stress (TADMUS) program to assess the accuracy and timeliness of team Tactical Decision Making (TDM) (Dwyer, 1992). ATPI first appeared in paper and pencil form and is based on Marshall's (1995) cognitive scheme model. It was built around communications between, and actions initiated by, Anti-Air Warfare (AAW) watch team members within a ship's Combat Information Center (CIC)

during the detect-to-engage sequence. Detect-to-engage training scenarios are pre-scripted and contain pre-programmed threats that the team must react to in a timely manner. Through ATPI, each detect-to-engage action (e.g., detect potential threat, issue first warning to potential threat) initiated by a trainee is scored by the trainer in terms of accuracy and timeliness (Dwyer, 1992). Thus, ATPI was designed to collect data in an air defense warfare training environment and may be used, in a modified form, within ATEAMS.

One drawback to the paper and pencil version of ATPI is that data entry and diagnosis must initially be done by hand, leading to a long delay between scenario completion and feedback. Feedback is generally more effective when delivered as close as possible to the end of a training exercise (Johnston, Smith-Jentsch, & Cannon-Bowers, 1997). In addition, the trainer is required to mark and discard paper forms, potentially missing critical events occurring during the scenario (generally, only one observation could be made per form) (H. Carroll, personnel communication, May 1, 2000). To compensate for these limitations, an automated ATPI application was created and placed on a handheld device. This application was developed for the Surface Combatant 21 Manning Affordability Initiative (SC21 MAI).

The ATPI application allows the trainer to move between a scenario screen, which displays all the critical radar tracks associated with a scenario and the scenario timeline (Figure 1), and a data collection screen (see Figure 2).
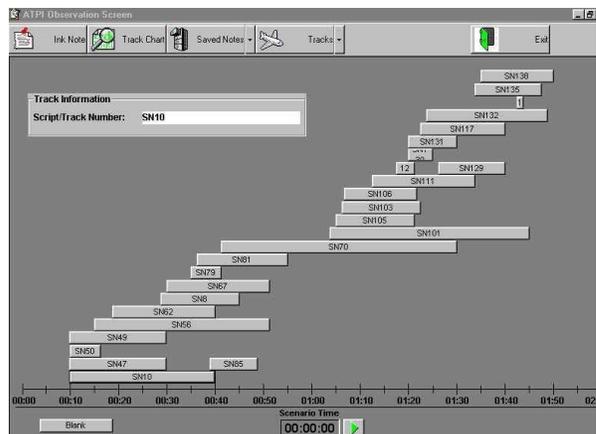

Figure 1. ATPI Scenario Screen.

To open the data collection screen, the trainer simply clicks one of the radar tracks (the gray rectangles seen in Figure 1). Within the data collection screen, all facets of the detect-to-engage sequence are listed, allowing the trainer the opportunity to score multiple AAW positions within the same timeframe. The

Figure 2. ATPI Data Collection Screen.

trainer scores a given action by selecting a trainee's position from a dropdown menu and rating the trainees action by clicking one of four buttons; correct planning or execution (Cp or Ce, respectively) or incorrect planning or execution (Xp or Xe, respectively). The trainer then saves the observation by clicking a button labeled submit. The observation is immediately time-stamped (J. Pharmer, personnel communication, May 1, 2000).

TDT
Team Dimensional Training (TDT) is a training methodology designed to aid instructors in training and evaluating teamwork skills. TDT provides a framework through which team members can develop shared mental models of teamwork. This is accomplished through a four step training cycle: briefing a team, observing a team's performance during a training exercise, diagnosing this performance, and de-briefing the team about its performance (Smith-Jentsch, Johnston, Cannon-Bowers, & Salas, 1997).

During the briefing phase the four teamwork dimensions delineated by TDT, and behaviors associated with each, are presented to the team by the trainer. The trainer also presents material pertinent to the given training exercise, e.g., rules of engagement (ROE) or maps. During the exercise itself, the trainer(s) attempts to gather positive and negative examples of behaviors that fall under each TDT dimension. One or two of the best examples (i.e., most relevant to the training objectives) are summarized in a post-training consensus meeting by the trainer(s). Then during the debriefing phase, the trainer attempts to facilitate discussion of the team's performance, providing positive and negative examples of team behavior under each dimension (only if the team could not come up with its own example, Smith-Jentsch, et al., 1997).

Like ATPI, TDT was originally a paper-based training methodology. All materials, including pre-briefing materials, TDT instructions, training exercise score sheets, summary tables, and debriefing materials were paper-based. One drawback to the paper and pencil version of TDT is that data entry must be done by hand. In addition, the instructor(s) must organize their paper-based scoring sheets following a training exercise, confer about their observations, diagnose performance problems, and prepare the debrief. The TDT booklet notes that, "For a group of instructors proficient at using TDT, this takes an average of 15 minutes" (p. 10). A group of less experienced trainers would obviously take longer, potentially leading to a long delay between scenario completion and feedback. The authors of TDT also recommend that one instructor be used for each of the four TDT dimensions (Smith-Jentsch, et al., 1997). In today's atmosphere of tight budgets, such a recommendation may not be possible to meet. To help counteract the aforementioned problems, a TDT application was created and placed on a HHD. This application was employed onboard the USS Constellation in 1996 as part of a demonstration of the BFTT system (D. Milanovitch, personnel communication, June 8, 2000).

One advantage of the TDT application is that it is entirely self-contained, meaning all materials germane to the methodology are placed on the application. Thus, the trainer(s) can use the HHD when briefing, observing, and debriefing a team. The trainer can present all briefing materials (e.g., ROE or maps) to the team via the HHD. Definitions and descriptions of each TDT dimension are also available to the trainer through the HHD, allowing less experienced trainers quick access to this information. During the training exercise itself, the trainer has access to a Scenario Worksheet screen, through which a variety of observations can be made (see Figure 3).
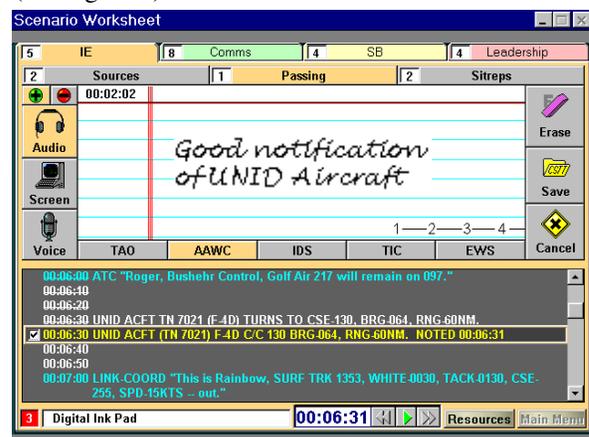

Figure 3. TDT Scenario Worksheet Screen.

Four color-coded tabbed areas are on this worksheet. Color-coded tabs allows the trainer to more easily delineate and access each TDT dimension and sub-dimension during the exercise. In addition, a scripted scenario scrolls and displays critical (and other) events to the trainer. The text within the scenario script is color-coded, with yellow text indicating critical events that can be observed. An embedded red checkbox located to the left of this text, when checked, indicates that the trainer observed the event. Future versions of MATE could have a checklist appear when the checkbox is clicked. The scrolling scenario text eliminates the need for the trainer to flip between pieces of paper, which only contain a portion of a scenario, when observing and scoring team performance. In addition, fast-forward and reverse buttons allow the trainer to move forward in the script of a scenario to see when the next critical event(s) are going to occur, allowing the trainer the opportunity to pre-plan and anticipate observations. The application's worksheet also provides the trainer with constant access to positive and negative scoring buttons as well a digital annotation device, audio capture, voice annotation, and the ability to digitally capture a trainee's display. The trainer's observations are automatically time stamped, in scenario time, eliminating the need for the trainer to separately record the time. The time stamp also provides a basis for automatically organizing the observations for post-exercise analysis. These features of the TDT application make it possible for a single instructor to make observations about multiple teamwork dimensions, eliminating the aforementioned recommendation of the use of multiple TDT trainers.

A debriefing diagnosis screen, titled Prepare Debrief, is also provided for the trainer (see Figure 4). This screen allows the trainer to view and organize all saved observations (e.g., digital ink notes, screen
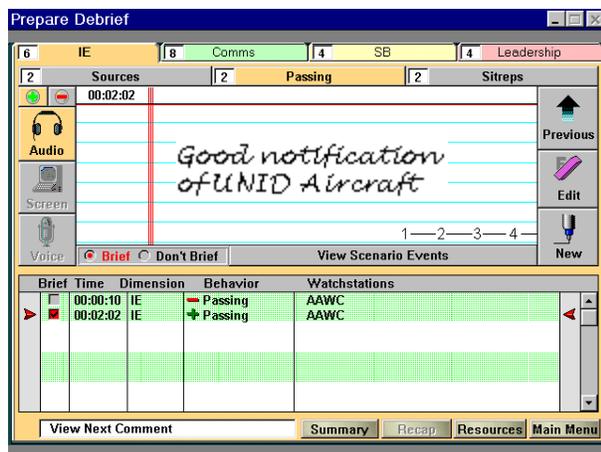

Figure 4. TDT Prepare Debrief Screen.

captures, and audio captures). For example, the digital ink notes are organized under each TDT dimension and sub-dimension and are sorted chronologically. A checkbox under the column labeled Brief (see Figure 4) is displayed next to each observation and provides a means by which the instructor can select the most pertinent observations, i.e., the observations that relate to the training objectives of the exercise. When checked, the observation will appear on the debriefing screen. The instructor can also view the pre-brief materials or the part of the scenario script that each observation is tied to.

The debriefing screen, titled Debrief the Team (see Figure 5), provides sample text on how to debrief each TDT dimension and sub-dimension, as well as example questions that can be read to help focus the team's discussion. As noted above, the instructor has access to all items previously selected from the
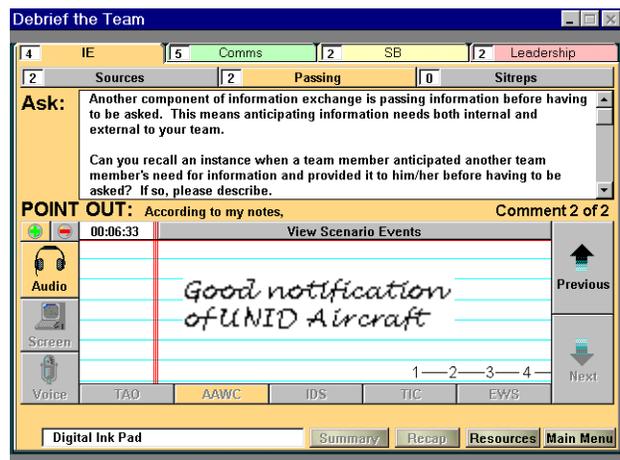

Figure 5. TDT Debrief the Team Screen

debriefing diagnosis screen, including the part of the scenario script that corresponds to the same observation, the pre-briefing materials, and all saved materials (e.g., digital ink notes or screen captures).

TRACE
The Tactically Relevant Assessment of Combat Events (TRACE) is a checklist-based application that was derived from the Targeted Acceptable Responses to Generated Events or Tasks (TARGETs) methodology and was adapted for use within the Division Artillery (DivArty) environment. TARGETs links events that occur during a training exercise to expected behaviors (Fowlkes, Lanes, Salas, Franz, & Oser, 1994). Initially paper-based, TRACE attempted to eliminate redundancy found on existing paper-based checklists that trainers use, finding more coherent groupings of the checklist items. The DivArty Mission Training Plan (MTP) was used to develop sixteen checklists that were then automated,

i.e., placed on a HHD (Burns, Gualtieri, Fowlkes, Reeves, Phipps, & Pierce, 1998). TRACE was then employed at a training exercise at Ft. Hood in January 1998. It was used to help trainers evaluate four elements of the DivArty staff, including Fire Support, Fire Control, Operations Center, and Targeting (J. Burns, personnel communication, June 6, 2000).

One of the benefits of automating performance measurements such as checklists is that a greater number of factors can be observed/rated. In addition, data can be collected, summarized, and analyzed more rapidly, allowing adaptive feedback to be given more quickly. Finally, personnel requirements, i.e., the number of trainers/observers needed for any given training exercise, can be reduced (Burns, et al., 1998).

Features found on the TRACE application include automated time stamping of all observations, digital ink, a wireless LAN that links raters (raters can send each other notes about the training exercise via the LAN), pop-up checklists, and on call checklists. Pop-up checklists are time-synced with scripted events within a training exercise while on call checklists can be brought up by the trainer at any time, e.g., if an unexpected event (and related behaviors) occurs during an exercise that the trainer wishes to rate (Burns, et al., 1998).

As noted, the pop–up checklists are time-synced to scripted scenario events. A scrolling script displayed on the HHD contains textual descriptions of each event, with each critical event within the script assigned a unique number. The scenario text is also color-coded, with yellow text indicating a critical event. As a critical event transpires the corresponding pop-up checklist, used to rate behaviors that should occur during that event, will appear. The trainer then rates the behaviors of the trainees by clicking Y, N, N/A, or N/O (Yes, No, Not Applicable, and Not Observed, respectively) next to each checklist item. If critical (or other) events are slightly out of sync with the anticipated (or observed) behavior, the trainer can minimize the pop-up or on call checklist by clicking a Hold button, maximizing and completing it later in the exercise. Figure 6 shows the TRACE data collection screen that contains the scrolling scenario script and a pop-up checklist that is linked with Event Four while Figure 7 displays an example of an On Call checklist.

TRACTs
The Tactically Relevant Assessment of Combat Teams (TRACTs) is also an adaptation of the
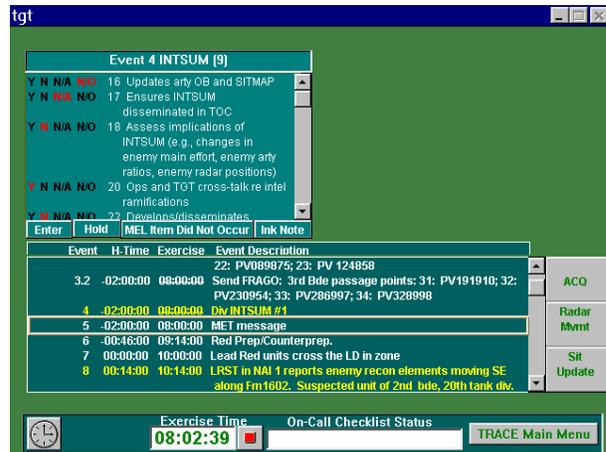

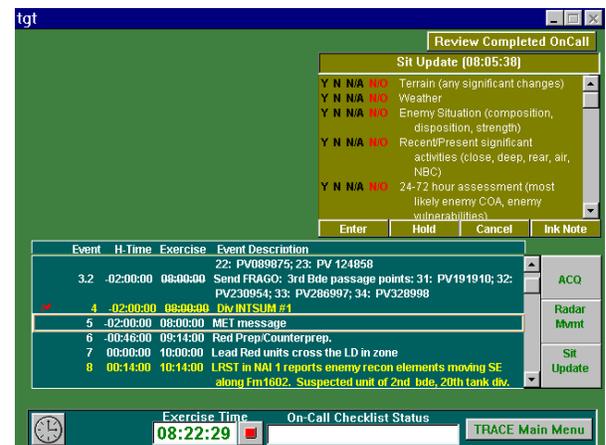**Figure 6.** TRACE Data Capture Screen with Event 4 Pop Up Checklist.

Figure 7. TRACE Data Capture Screen with an example of an On Call Checklist.

TARGETs performance methodology. TRACTs was designed for use within the simulation-based Army artillery, or DivArty, environment and was used as part of the Crusader Concept Experimentation Program (CEP) at Ft. Hood in the Fall of 1998. The goal of the program in which TRACTs was employed was to assess the impact of different tactics, techniques, and procedures (TTP) on human performance. Different levels of the Crusader team were assessed, e.g., the Brigade Fire Support Element and the Platoon Operations Center (Fowlkes, Dwyer, Milham, Burns, & Pierce, 1999; Stretton, Burns, Phipps, & Reeves, 1999).

TRACTs is a follow-on application that utilized concepts developed for TRACE. Unlike TRACE, TRACTs uses a Windows-like menu structure through which the trainer can select checklists (see Figure 8). TRACTs also allows real-time editing of TRACTs also allows real-time editing of saved checklists, digital ink notes, or voice notes (see Figure 9).
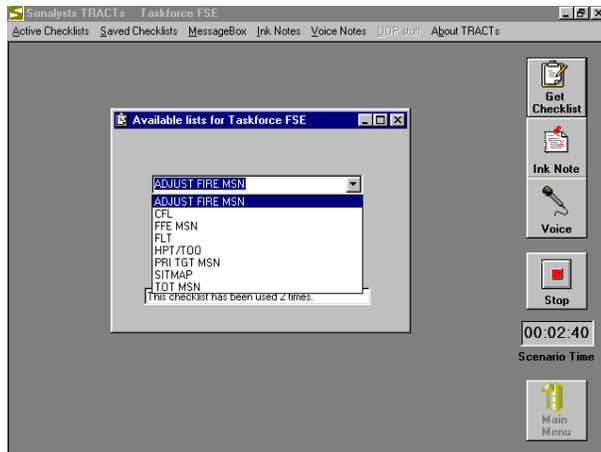
Figure 8. Clicking the "Get Checklist" Command Button opens the TRACTs Checklist Window.
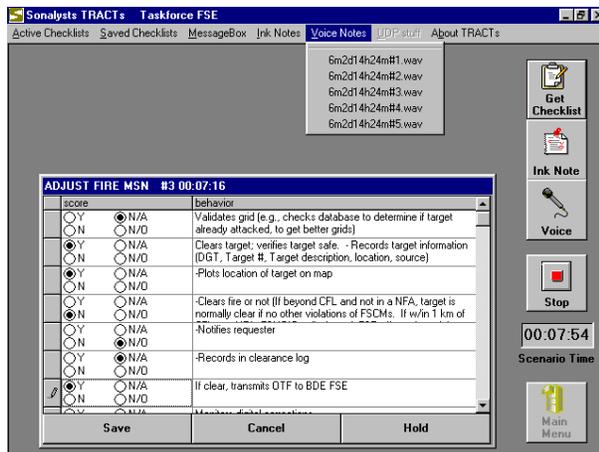

Figure 9. TRACTs "Adjust Fire MSN" Checklist. Note the Drop Down Window (of saved voice notes)

TRACTs is similar to TRACE in that it employs On Call checklists to be used by the trainer(s) for rating performance (the checklists were generated from existing Standard Operating Procedures). A text box displays the number of times the called checklist has been saved. With the use of a radio-link, instructors can send each other notes (either digital ink or voice) via the TRACTs interface. However, TRACTs does not employ a scrolling scripted scenario window. Therefore, TRACTs was used for scoring events that occur spontaneously (i.e., events that were not scripted) during a training exercise. TRACE, on the other hand, was used for scoring both unscripted and scripted events (Fowlkes, et al., 1999).

AETS
The Advanced Embedded Training (AET) System, or AETS, is a multilayered training platform that provides real-time and automated online feedback, team performance measurement and diagnosis, and computer-assisted coaching. The types of information

that can be captured through AETS include eyetracking data, keystrokes, voice (through speech recognition), and instructor input (via MATE). AETS was demonstrated at the AEGIS Training and Readiness Center (ATRC) in Dahlgren (D. McDonald, personnel communication, June 7, 2000). The expectation is that AETS will be transferred to the AEGIS Combat Training (ACTS) and Battle Force Tactical Training (BFTT) programs through ATEAMS.

The training method loaded on MATE, used by the AET instructor, is TDT and has all the features listed under TDT above, e.g., prepare and present prebrief or debrief materials. In addition to those features, MATE is the instructor's interface to the simulation and intelligent tutoring systems of AETS. Before the training exercise begins, the instructor can configure the real-time feedback that the trainee (or team) will receive through a Feedback Configuration option. The options include providing automated, performance based feedback that is timed with scenario events as well as allowing the instructor to generate and send feedback messages to the trainee (see Figure 10). During the training exercise, the instructor is alerted to a trainee's performance and is
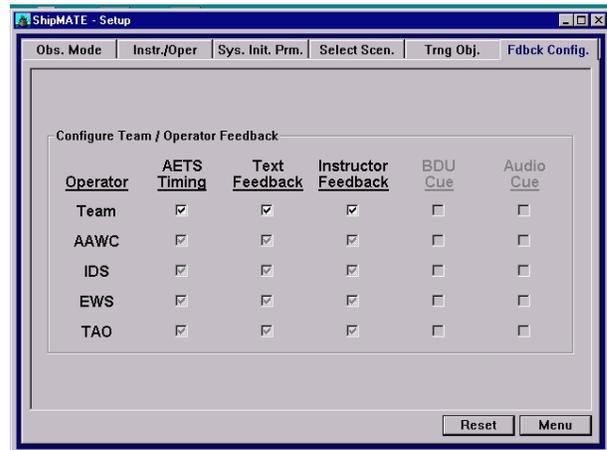

Figure 10. AETS Feedback Configuration Screen.

given several suggested canned feedback messages that can be sent to the trainee. Once selected, the instructor can send the message to the AETS feedback display by clicking the Send button (see Figure 11). The trainee will not be able to tell if the feedback message came from the intelligent tutoring system or the instructor. Although not implemented for the Dahlgren demonstration, instructors can send each other messages through the HHD.

Like the TDT application above, the AETS MATE application can digitally capture the trainees display screen. However, AETS has the capacity to digitally record both audio and video
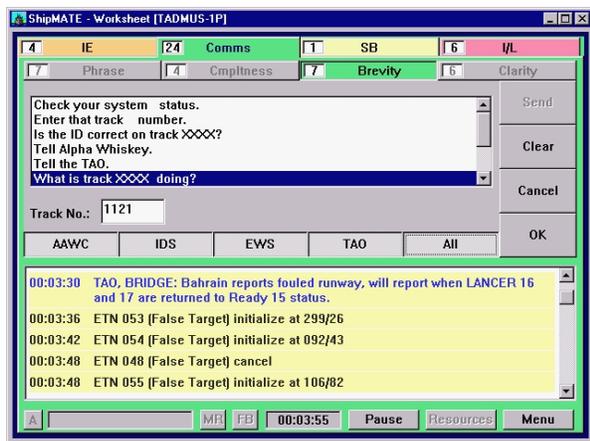
Figure 11. Instructor Feedback Messages (upper text).

information. Therefore, the AETS MATE application can provide the trainer with a link to a video replay of a trainee's performance that the trainer can access during the debriefing session.

FIELD TRIAL OBSERVATIONS

As noted above, each application has been tested and/or demonstrated in the field. What follows is a summary of observations made during the field trials for each application.

ATPI

Three developers of the ATPI application, one a Human Factors specialist, the others Subject Matter Experts (SMEs) were asked to give their opinion about the utility of the application compared to the paper-based version of ATPI. All three strongly emphasized that the most important feature of the ATPI application is its ability to time-stamp observations, i.e., ATPI allowed them to get accurate time stamps of the behaviors/actions being recorded. The Human Factors specialist noted that the paper-based ATPI could yield over 100 individual papers, so the application was "ideal" for post-data analysis (i.e., easier than typing in data from 100 papers). The SMEs also felt that the application allowed for more detailed data to be recorded.

One problem reported by the SMEs, not directly related to the application, was the size of the HHD screen, i.e., they wanted the ability to view the scenario and data collection screens simultaneously. In addition, data were lost because the HHD battery ran dry. One SME felt that using the HHD's pen took a bit of getting use to. Finally, both SMEs felt that ATPI needed an undo button. That is, during a training exercise the instructor might have saved an observation, only to realize later that a mistake was made. However, the instructor had no method that allowed him to call up the observation to delete (or

change) it. As a result, some of their data had to be scrubbed because erroneous observations were saved.

AETS (TDT)

As mentioned above, AETS was demonstrated at ATRC in Dahlgren. Three instructors taking part in the exercise were asked to fill in a questionnaire after each scenario run (six-point likert scale; 1 = strongly disagree, 6 = strongly agree). Six questions related directly to MATE. The questions, along with the resulting means, can be found in Table 1.

Excerpt from AETS/MATE Capabilities Questionnaire	Mean Rating
The MATE user interface was easy to use.	4.67
The button labels were understandable.	5.67
The system was responsive when you clicked on a button.	5.33
The functions that were available make sense to you.	5.67
The HHD was easy to hold and use.	4.00
All the words on the display were readable.	5.67

Table 1. Excerpts from the AETS/MATE Capabilities Questionnaire.

Since the data in Table 1 are descriptive and are based on a small sample, no definitive conclusions can be made. It would appear that the TDT interface generally had good utility, especially the button labels, and that text resolution was acceptable.

TDT

One SME was asked to compare the TDT application with the paper-based version of TDT. He noted that he could make more observations with the application (e.g., 40-50 with the application compared to 11-12 with paper) but is not sure if quantity = quality, i.e., were those extra observations really useful or important? He also felt that the most useful features of the application included the ability to rate behaviors with the plus/minus buttons found on the worksheet, and correlating his observations with a specific point in the scenario. He also felt that the prepare debrief and debriefing screens were useful, helping him to address the team about their performance. He felt this would lead to a realistic assessment of their behavior and help find areas for improvement.

TRACE

One developer, familiar with both TRACE and TRACTS, was asked to give his opinion about the TRACE application. He felt that the most useful feature was the pop-up checklists.

During a field test, Burns, et al. (1998) collected data from the paper-based and application-based versions of TRACE. Four users were employed, two of which used both versions of TRACE. Burns, et al. (1998) report that the users felt that the TRACE application was easy to use and made measurements more controllable. They also note that users liked the fast forward tool that allowed them to scroll forwards in the scenario to preview upcoming scripted events, then being able to immediately return to the current event (by clicking the Resume Auto Scroll button). Users also liked the digital ink tool as well as the automatic generation of the event based (i.e., pop-up) checklists. The users also indicated the need for a virtual keyboard, the ability to recall and edit digital ink notes, and the ability to record a digital ink note that was not linked to a checklist.

TRACTs
One developer, familiar with both TRACE and TRACTS, was asked to give his opinion about the TRACTs application. He felt that the most useful feature was TRACT's ability to quickly convert data into a presentable form, e.g., into PowerPoint slides with charts, graphs, etc.

Fowlkes, et al. (1999) report that during the Crusader effort, data collectors frequently used the wireless LAN capability of TRACTs to communicate with each other. This allowed the instructors to coordinate and exchange critical team performance data. Unfortunately, a problem with the wireless LAN required final downloading of data to be done through a wire connection to a PC.

In the Crusader exercise, they also used the digital ink note-taking tool to record comments and observations. Fowlkes, et al. (1999) report that these notes were useful when trying to interpret the performance data obtained from the training exercise. They also note that quick data analysis was possible through the TRACTs application, something that is not possible with a paper-based data collection tool.

In summary, the MATE applications generally reduced workload, especially when the original performance measurement tool was linked to paper-based checklists. They also reduce the workload associated with post-training data entry and analysis. They may help reduce the complexity of tracking multiple person teams. They allow trainers the ability to communicate with each other during a training exercise. They provide trainers with the ability to collect and store data in real or near-real time. Finally, they provide a way for the trainer to accurately time stamp observations and quickly rate behavior.

TAXONOMY OF TOOLS

We have described several variations of the original ShipMATE hand-held performance measurement system, currently referred to as MATE. Each application contains a suite of tools that have been customized (grouped) to fit a variety of training environments. A listing of these tools has been compiled in Table 2 along with their associated purpose and/or function. The table is divided into the four phases that typically occur during training: a pre-training briefing, the training exercise (or scenario) itself, diagnosis of trainee(s) performance, and a post-training debriefing. Each tool is listed under the phase that they are most likely to be used. Finally if an application has a given tool, its abbreviation is listed under the applications column.

SUMMARY AND GUIDELINES FOR FUTURE DEVELOPMENTS

Researchers at the Naval Air Warfare Center Training Systems Division (NAWCTSD) have developed multiple instructor aides for performance measurement hosted on hand-held computers. This paper has reviewed the multiple tools that have been developed for these aids and presented a taxonomy of suggested utilization of the optional screen functions. Since the development of the original prototype, i.e., ShipMATE, additional training and evaluation software applications have been developed and demonstrated for Air, Surface, Ground, and Sub-Surface domains in both classroom and operational settings. These software applications were developed by selecting tools listed in Table 2, each tailored to accommodate various performance measurement methodologies including outcome and process measures aimed at capturing individual and team performance.

The viability of HHD technology for operational settings has been demonstrated by several diverse field trials, including sea trials aboard AEGIS ships (ATPI, TDT), field operations during Army exercises (TRACTs, TRACE), classroom training at the Aegis Training and Readiness Center (ATRC) (AET, TDT), airborne data collections of aircrew teams (TDT and the Reconfigurable Aid for Training and Evaluation (RATE, the latter still in development)), and training conducted during I/ITSEC 99 "Operation Desert Rats" distributed simulation (TDT). Table 3 lists each application, the funding source, operational community, target training audience and the date the application was used.

The continuing emphasis on decreased manpower and increased automation in tomorrow's training systems, coupled with the increasing capacity of

Tool	Purpose/Function	Applications
Pre-Training Brief		
Pre-Briefing Materials	Allows the trainer to access/display all briefing materials from one screen, e.g., rules of engagement, maps, routes to use, etc.	AETS, TDT
Scenario Settings Screen	Allows the trainer to select various scenario options, e.g., scenario speed, time to start scenario, scenario type, observation method, team member assignment, etc.	ATPI, AETS, TDT, TRACE, TRACTs
Training Exercise (Scenario)		
Audio Capture	Allows the trainer to capture audio from the trainees. If linked to a digital recording system, allows the trainer to "frame" the audio timeline that he wishes to capture.	AETS, ATPI, TDT, TRACTs
Color Coded Text	In scripted scenarios, the text of critical events can be color coded (see scrolling scripted scenario below).	AETS, TDT, TRACE
Digital Ink	Allows the trainer to write and save digital notes relating to training incidents/behaviors.	AETS, ATPI, TDT, TRACE, TRACTs
Drop Down Windows	Allows the trainer to quickly access data, e.g., track numbers, watch station assignment, checklists, etc.	ATPI, TRACTs
Editing	Allows the trainer to edit digital notes.	ATPI, TRACTs
On Call Checklists	Provide the trainer with a means of scoring trainees' behavior when unexpected events occur during an exercise.	TRACE, TRACTs
Pop-up Checklists	Pop-up checklists display during critical events in the training exercise. Checklist items are tied to expected behaviors.	TRACE
Scenario Clock	Displays the scenario time that is synced with scripted events/pop-up checklists. Can be set to standard, scenario, or Zulu time.	AETS, ATPI, TDT, TRACE
Scenario Controls	Fast forward, reverse, or pause a scenario script.	TDT
Screen Capture	Allows the trainer to capture a display that a trainee is viewing, e.g., a radar display. Digital ink can be used for notation on the capture.	AETS, TDT
Scrolling Scenario	Displays textual descriptions of critical (and other) events. The text can be color-coded to delineate the criticality of the event.	AETS, TDT, TRACE
Scenario Speed	Allows the trainer to select the speed at which as scenario is played.	TDT, TRACE
Scenario Timeline	Allows the trainer to view an entire scenario at once.	ATPI
Time Stamping	Time stamps input (e.g., digital ink) made by the trainer.	AETS, ATPI, TDT, TRACE, TRACTs
Voice Annotation	Allows the trainer to make notes by capturing his/her own voice.	AETS, TDT, TRACTs
Diagnosis of Trainee Performance		
Debrief Organization	An organization screen through which the trainer can select the items he wishes to include when debriefing the trainee(s).	AETS, TDT
Post-Training Brief		
Debrief Presentation	Provides a method through which the trainer can present the debriefing materials to the trainee(s).	AETS, TDT

Table 2. Taxonomy of ShipMATE Tools, Descriptions of their Purpose/Function, and Associated Applications.

hand-held computer and palm pilot technology is creating endless opportunities for hand-held performance measurement devices. However, technology is only part of the solution. The main ingredient in determining the utility of the device will be the effectiveness of the underlying performance measurement methodology.

The methods that have been used by NAWCTSD to develop the applications described in this paper include:

1. Survey of current practices. Are there measures of performance or measures of effectiveness (MOPs/MOEs) that are accepted by the training community?

2. Identify the individual vs. team, and outcome vs. process, goals of the training. Are there proven, documented methodologies (possibly paper-based) that can be implemented?

Application	Funded by	Target Group	Operational/Experimental Setting	Date Employed
AETS (TDT)	Office of Naval Research	Navy Surface - CIC	Demonstrated at ATRC in Dahlgren.	Summer 98 and 99
ATPI	Surface Combatant - 21 Manning Affordability Initiative	Navy Surface - CIC	Onboard Ship - DDG and CG Class.	Fall 99 and Summer 00
TDT	Office of Naval Research	Navy Surface – CIC	USS Constellation: Demonstration of BFTT. I/ITSEC 99 MMTT Demo.	Summer 96 Winter 99
TRACE	Army Research Lab (Ft. Sill)	Army (DivArty) – Four Elements (TGT, FCE, OPS, & FSE)	Ft. Hood.	January 1998
TRACTs	Army Research Lab (Ft. Sill)	Army – Various Tactical Operation Centers associated with Army artillery	Ft. Hood - Crusader Concept Experimentation Program (CEP).	Fall 1998

Table 3. Application Domains**.**

3. Identify all sources of performance data, such as voice and video recording through a simulation system. The goal of the HHD application is to be the instructor's interface to the simulation system. Providing the ability to "tag" the desired data (such as a voice clip) for quick retrieval and use in a debrief.

4. Select and tailor the tools listed in table 2 that are necessary to accomplish the specific training goals as identified in step 2, above.

5. Continually call upon subject matter experts, and the end-users of the application for their recommendations, reviews and guidance. Collect and implement feedback from beta testing with the users. Several iterations may be required to achieve an optimum design.

The applications summarized in this paper represent just a few of the many possibilities for hand-held performance measurement devices. We offer these tools and practices as guidelines for developing hand-held instructor aides for future training systems.

REFERENCES

Burns, J., Gualtieri, J., Fowlkes, J., Reeves, D., Phipps, D., & Pierce, L. (1998, October). Dynamic assessment: Using technology to enable and extend performance measurement. Poster session at the Human Factors and Ergonomics Society Annual Meeting, Chicago, IL.

Burns, J. J., Stretton, M., Phipps, D., & Johnston, J. H. (1999). Developing a common user interface for managing advanced embedded training. *Unpublished technical report*. Orlando, FL: Naval Air Warfare Center Training Systems Division.

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

Fowlkes, J. E., Lane, N. E., Salas, E., Franz, T., & Oser, R. (1994). Improving the measurement of team performance: The TARGETs methodology. *Military Psychology, 6*, 47-61.

Fowlkes, J. E., Dwyer, D. J., Milham, L. M., Burns, J. J., & Pierce, L. G. (1999). Team skills assessment: A test and evaluation component for emerging weapon systems. *Proceedings of the 1999 I/ITSEC* (pp. 994-1004) [CD-ROM]. Arlington, VA. National Training Systems Association.

Johnston, J. H., Smith-Jentsch, K. A., & Cannon-Bowers, J. A. (1997). Performance measurement tools for enhancing team decision-making training. In M. T. Brannick, E. Salas, & C. Prince (Eds.) *Team Performance Assessment and Measurement: Theory, Methods, and Applications.* Hillside, NJ: LEA.

Naval Air Warfare Center Training Systems Division. (1999). *TDT: The handbook for team dimensional training* [Brochure]. Orlando, FL: Naval Air Warfare Center Training Systems Division.

Pruitt, J. S., Burns, J. J., Wetteland, C. R., & Dumestre, T. L. (1997). ShipMATE: Shipboard mobile aid for training and evaluation. *Proceedings of the 41st Annual Meeting of the Human Factors and Ergonomics Society*, 1113-1117.

Smith-Jentsch, K. A., Johnston, J. H., Cannon-Bowers, J. A., & Salas, E. (1997). Team dimensional training: A methodology for enhanced shipboard training. *Proceedings of the 19th Annual I/ITSEC Conference* [CD-ROM].

Stretton, M. L., Burns, J. J., Phipps, D., & Reeves, D. (April, 1999). Training management module and hand-held device performance measurement: Year 1 report. *Unpublished Technical Report*. Orlando, FL: Naval Air Warfare Center Training Systems Division