Human Factors in Simulation and Training

Distributed After Action Review for Simulation-Based Training

By Emily E. Wiese, Jared Freeman, William J. Salter, Emily M. Stelzer, Cullen Jackson
Aptima, Inc.

Introduction

In this introduction, we begin by defining "distributed after action review (AAR) for simulation-based training" and deriving a few implications from the definition. We then touch on several issues that seem general across distributed AARs for simulation-based training, although we note that the diversity of such training means that exceptions might well be found. Finally, we briefly outline the balance of this chapter.

Definitions

"Distributed" means geographically dispersed, which can range from being in separate rooms to thousands of miles of separation. An "After-Action Review" (also "AAR," "debrief," or "debriefing") is a facilitated discussion of training performance in which the basic goal is to enhance subsequent trainee performance[1], generally conducted soon after the training event. (We discuss AAR functions in more detail in the next section.) It is important to note that by "distributed AAR" we mean that the AAR itself (and, typically, the instructors) will be distributed, not only the learners. That is a vital constraint, since it means that instructors cannot all observe the same aspects of training or meet face-to-face to discuss what happened. They must use collaborative technology of some sort, which can range from simple teleconferences to sophisticated computer systems.

"Simulation-based training" means that learners use simulators, rather than engaging in live exercises. Such simulators are typically highly technological and range from simple desk-top flight simulators to submarine control rooms identical to real ones except for the lack of a surrounding submarine and ocean. For the purposes of this chapter, we assume that the simulators involve computers and that the computers collect data on what is occurring during the exercise and how trainees respond to events during the exercise. Of course, the quantity and utility of such data can vary widely, and this can greatly affect the detail and coverage provided in the debrief.

[1] AARs are frequently conducted for actual missions as well. These can be very valuable for enhancing tactics, gathering intell, and improving procedures and performance, but they are not the subject of this chapter.

AARs for distributed simulation-based training are growing in importance largely because distributed simulation-based training is increasing in prevalence. This increase is in turn driven by several factors. First, more missions are coalition and/or joint, and the heterogeneous participants in such missions tend to be more geographically distributed, and thus harder and more expensive to coordinate for co-located training, than participants from one service. This in part reflects the broadening of missions beyond those traditionally handled by the military. Note that the broadening of missions and their joint/coalition emphasis also tend to require that a wide range of skills be trained, specifically including teamwork and coordination skills, and implies that one overall (often called "community" or “mass”) debriefing may not address all training requirements. We will return to this point.

Second, cost pressures more broadly are making it harder to support the (often considerable) expense of bringing people together for the purposes of simulation-based training. And third, rapid advances in enabling technologies make it relatively easy to use existing simulators for distributed simulation. High levels of network reliability, bandwidth, and speed are dropping in price far more quickly than new simulation technologies are being introduced, making the basic technology infrastructure needed for distributed simulations more available. In particular, DIS (distributed interactive simulation) and HLA (high level architecture) – the two communications standards used for almost all military simulations – can run over wide-area networks as long as those networks have adequate performance characteristics, which means that simulators that often had to be in adjacent rooms to interoperate effectively can now be separated by hundreds or thousands of miles.

Distributed simulation-based training will involve multiple simulators. And typically, multiple platforms, weapons systems, and/or elements will be involved. (This is not necessarily true – for example, a strike lead and a wingman could participate in a coordinated mission using simulators in different locations – but this is the usual circumstance.) And there will typically be at least an implicit hierarchy of training objectives, some for the individual elements, some across elements, and some for the mission as a whole.[2] Debriefings for distributed simulation-based training generally reflect two levels of the basic hierarchical structure of objectives: separate debriefs for each element or platform and a community or mass AAR that addresses coordination across elements and mission objectives.

Some Issues in Distributed AARs for Simulation-Based Training

Different information is required for each element or platform and for the community debrief, and different expertise is generally required to deliver the debriefs. This creates several complexities. First, it is desirable that **an instructor** (or a designated participant-

---

[2] Designing and implementing scenarios for distributed simulation can be more complex for more specialized training. And the AAR obviously has crucial dependencies on the scenario. However, the important topics of training objective development and scenarios design are beyond the scope of this chapter.

instructor) **be physically present for each separate "cell" or team of the distributed simulation**. Since some important actions (or lack of actions) may not be recorded in sufficient detail for immediate post-exercise analysis by automated components of the distributed simulation, humans can provide a level of detail that can be quite valuable. For example, if participants communicate by voice, even if all voice communications are recorded, language processing software is not currently adequate to support pedagogically useful analysis. Furthermore, replay of *all* communications for post-exercise analysis requires too much time and must be synchronized with replay during the review in order to be meaningfully linked with other actions.

Second, identifying cross-platform issues to be addressed in the community debrief, and deciding how to characterize those issues, requires **information fusion across the distributed simulation platforms**. By definition, those issues will involve interaction, coordination, and/or communication across platforms or elements. The identification of problems of that sort can be difficult, and the diagnosis and suggested remediation for problems will tend to be more so. This fusion must be addressed via procedures, technology, and preparation of the instructors.

Thus, more elaborate debrief preparation is generally needed, both because instructors must gain insight into events and actions that they did not observe and because issues to cover in debriefs must be divided among elements or platforms and between element debriefs and the community debrief. We believe that **more formal and more extensive use of computer-supported debrief preparation tools is needed** to effectively conduct AARs for distributed simulation-based training than for traditional co-located training. This is true to a lesser but still significant extent even if a single element or platform is being trained, since distributed instructors must share insights and build up reasonably extensive shared situation awareness.

Computer-supported methods of rapidly moving through and analyzing performance can be of particular value in distributed training, since more cognitive processing will tend to be required by distributed instructors than in co-located training. Importantly, while fully shared situation awareness may *not* typically be possible, even for distributed training of a single element, it is needed to understand the complete mission timeline and the ways in which trainee performance propagate across time and elements. Well-designed computer-supported AAR tools should assist in sharing the debriefing workload across instructors, allowing each to address what he or she has observed in detail, thus facilitating a more global awareness of performance across the exercise participants..

Ideally, these analysis and debriefing technologies will also function as **collaboration tools for distributed instructors**, since they operate under considerable time and task pressure. The time pressure is obvious – AARs typically take place within an hour of completing the training exercise, while experiences are fresh in participants' and instructors' minds. Task pressure comes from the fact that failing to address important aspects of the training in the AAR can result in seriously impaired learning, which, of course, attenuates the entire purpose of the exercise.

The Rest of this Chapter

In the next section we discuss *traditional debriefing techniques* in some detail. We outline the goals and functions of AARs and characterize the strength and weaknesses of the methods employed. We then go into more detail on the *challenges and opportunities* afforded by the increasing importance of distributed AARs. We operationalize the previous discussion by addressing the *requirements* for distributed AARs, and then discuss *current methods* for conducting them. Finally, we conclude with a more speculative discussion of *the future of distributed AARs*.

Functions and Current Methods of AAR

AARs fulfill diagnostic, instructional, and social functions. We define these functions here, then turn to a brief review of the current After Action Review techniques and technologies that attempt to support them.

Functions of AAR

The principle function of AARs is instructional: it must convey the right lesson to the right people at the right time. In more mechanistic, cognitive terms, an AAR must help the learner recall from the recent training specific episodes of *correct* performance in context, and generalize this knowledge so that it can drive performance in similar (but not identical) future circumstances. An AAR must also help the learner recall specific instances of *incorrect* performance, discourage repetition of that performance in the future, cue recall of correct performance knowledge, and associate it with similar circumstances. These activities should be targeted only at trainees who need to learn and who have the capacity and motivation to do so.

An AAR is not just a process for delivering lessons, it is also helps participants to discover those lessons. That is, after action reviews help trainers and trainees discover performance failures (and successes) and diagnose their causes.

Finally, AARs serve a social function. For all involved, they are an opportunity to show and assess technical competence – the ability to discriminate good from bad performance and diagnose its cause – and social competence – the ability to convey critiques of oneself with candor, and of co-participants or trainees with diplomacy.

In sum, AARs support (D1) diagnosis of performance, (I1) recall of performance in training, (I2) understanding of expert performance, (I3) generalization to future situations, and (S1) assessment and display of competence.

Methods of AAR

Techniques and technologies for AARs have evolved to support these functions more or less completely, more or less well.

Simulators typically support the instructional function of AAR through replay. Platform simulators, such as flight simulators, and large-unit simulator systems, such as Joint Semi-Automated Forces (JSAF), commonly record the course of selected simulation events, enable trainers to tag or bookmark instructive moments in that event stream, and replay those events using standard video controls (play, fast forward, etc.). Serial replay should have strong effects on learning. It conforms to the serial structure of the episode in memory and thus serial replay should ease recall (I1). It should reinforce memory for normative sequences of events (e.g., mission phases) (Schank, 1982) and thus it should help trainees to generalize from the specific episode to its class (I3). Finally, serial replay should help trainees recognize actions that snowball into disaster (or success) as a scenario evolves and thus serial replay should support the diagnosis function (D1) of AAR and the skills underlying diagnosis and, one hopes, prognosis.

The devil is in the details, however. Replay is often implemented in ways that help trainees to recall and learn the whole but not the parts, the progress of their unit but not their vehicle, and the flow of the mission but not its parts. For example, replay is often implemented as a set of icons moving over a tactical map (see Figure 1). This is a useful representation for cueing recall of the tactical state of forces and tactical actions of units. However, many simulators don't help individuals recall the situation by, for example, presenting the scene as a pilot, tank driver, or other operator *perceived* it. Replay systems rarely record and represent operators' displays and instruments, nor do they support recall of *responses* because the systems generally do not record and represent the trainee's use of controls. Still fewer simulators will record all or some of these viewpoints. Rare exceptions, mainly in the field of aviation training, include an F-16 distributed debriefing system developed by the Air Force Research Laboratory (AFRL) that combines a central tactical view with instrument displays on each side (Figure 2) (Sidor, 2002). Similarly, the Dismounted Infantry Virtual After Action Review System (DIVAARS, developed by the Army Research Institute and the University of Central Florida) provides multiple viewpoints of the simulation space during replay. These systems encourage operators to have detailed discussions about how they perceived the environment and the actions they took in it.

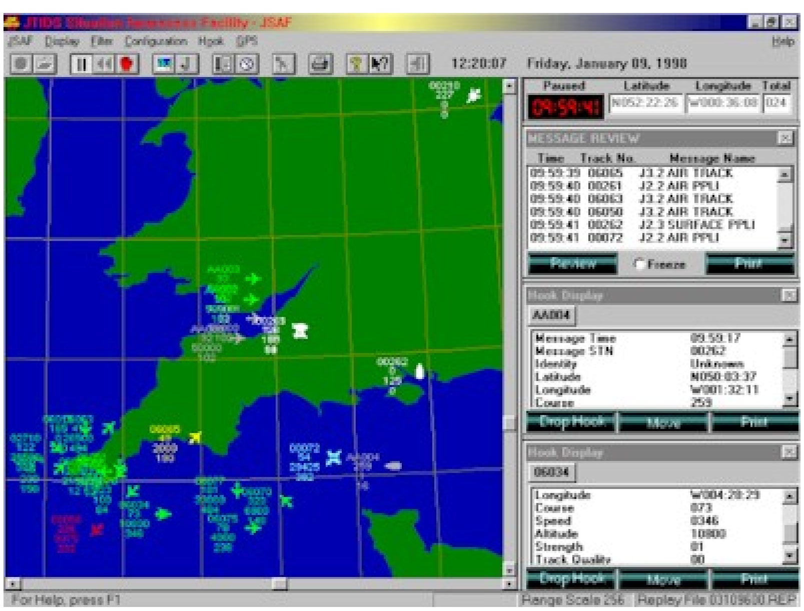

Figure 1. A typical tactical replay (left) shows the movement of forces in replay and associated Link 16 messages. (Source: http://www.tadil.net/netscape/images/jsaf.jpg.)

Figure 2. The Air Force Research Institute developed a tactical display coupled with an instrument replay supports recall of a pilot's situation and performance (left) (Source: http://www.afrlhorizons.com/Briefs/Sept02/HE0205.html).

Above, we noted that replay may help trainees learn the course of an overall mission, but not its parts. If an AAR is to help trainees generalize parts of the current scenario to future scenarios (I3, above), it may be useful to show them multiple instances of a class of situation and of response. Astute trainers can use the bookmark features of simulator AARs to tag similar situations and responses. In our experience, this is rarely done, however. AARs more often proceed from the first scenario to the last, rather than jumping between like events. AAR technology rarely helps trainers identify and navigate from one instance of a class of events to the next because simulation systems are rarely instrumented with measurement systems that can categorize events and human responses to them. One exception (Salter, et al., 2005; Figure 3) categorizes events, assesses performance in those events, displays those categories and assessments, and links each instance directly to its replay. This design gives trainers random access to multiple instances of a given class of events and thus supports trainees as they attempt to generalize from instances to the larger class.

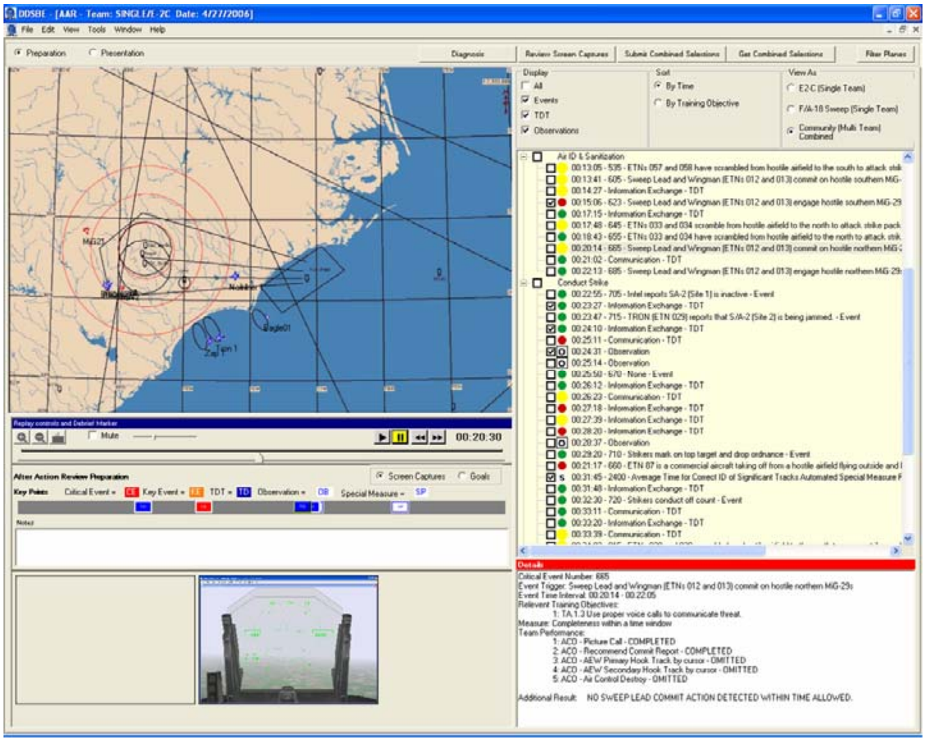

Figure 3. Assessments of categories of performance (right) are each automatically indexed (bookmarked) to the replay (left) (Salter, et al., 2005).

If AARs are to help trainees learn expert alternatives to their incorrect performance (I2, above), it must present them in some way. Few, if any, high end simulators (e.g., flight simulators and driving simulators with physical cockpits) are capable of generating examples of expert performance because they do not incorporate computational models of expert behavior. Thus, AARs raise alternative behaviors to the attention of trainees largely through discussion; alternatives are said, not shown. This may be sufficient in some circumstances. However, recent empirical research (Shebilske, et al., in press) demonstrates that describing expert solutions to complex problems (e.g., team planning and execution of military air missions) produces learning outcomes that are reliably inferior to describing and "playing" solutions using visual animations. In this research, expert solutions were generated by an optimization model. In traditional intelligent tutoring systems, these solutions are generated by heuristic or rule-based models of expertise.

The diagnostic function of AARs typically is supported by debriefing techniques, not technology. These techniques encourage trainers and trainees to identify and analyze strengths and shortcomings in performance. The Navy and the Air Force, for example, decompose AARs for large simulated and live exercises into independent debriefs of small elements or packages and overall debriefs involving the entire training audience (community or mass debriefs). The element debriefs typically identify specific performance failures that support diagnosis in the community debrief. The Army has codified its method in a set of questions that soldiers explore in AARs: What was supposed to happen? What happened? What accounts for the difference? (Dixon, 2000). A set of AAR guidelines (e.g., "Call it like you see it," "No thin skins") encourages participants to think critically and to be candid in their review of events. Similarly, the commercial aviation community espouses AAR methods that engage trainees in identifying and diagnosing performance failures. Studies of these methods and their impact on diagnostic quality are rare. However, one analysis of AARs in commercial aviation (Dismukes, et al., 2000) found that instructors often failed to engage trainees in diagnostic (or any) discussions during AAR, and instead dominated these sessions with monologues concerning their own observations. When techniques fail in this way, trainers and trainees cannot fall back on diagnostic technologies. AAR systems are generally incapable of generating diagnoses because they do not incorporate expert behavioral models against which to compare trainee performance. Nor do most AAR systems record data concerning trainee performance or compute measures that summarize that performance, attribute effects to individual performers, or relate causes to effects.

Finally, AAR techniques (but not the technologies) support the social processes (S1, above) by which participants assess the competence of their colleagues and assert their own. The Army AAR procedures encourage participation, as do the commercial aviation techniques mentioned above. However, current AAR technologies do not provide any features that support assessment and assertion of competence. Technologies that might provide this support include polling and group decision support tools that elicit or require input from all participants, and AAR authoring tools with which participants could focus attention on key events. The lack of tool support is particularly troublesome in distributed AARs, where reliance on teleconferencing may even hinder this social function by anonymizing participants or excluding some trainees from full participation due to technology difficulties (e.g., poor audio quality).

The functions of AAR – instructional, diagnostic, and social – are partially supported by AAR technologies. AAR techniques help trainers to address some of the remaining functions, particularly diagnostic and social functions. However, trainers have a spotty record of applying these AAR techniques well, and the advent of distributed debriefings may make it even more difficult for them to use these methods, as we discuss below.

Challenges of Distributed AARs

Recent advances in technology have allowed for the development of coordinated simulation tools, which can be used to simultaneously train groups of individuals who are dispersed across several geographical locations. As outlined above, these distributed training exercises can be conducted with less cost and risk than traditional live training events, allowing diverse groups of individuals to collaboratively train whole mission exercises more frequently than was ever before possible (e.g., Dwyer, Fowlkes, Oser, & Salas, 1996). While distributed training exercises can produce more effective and routine training events, this new approach to training can complicate debriefs that follow those events. We discuss the potential challenges that distributed training can pose to debriefing within the framework of the five key functions of an AAR outlined above, namely diagnosis of performance (D1), recall of performance in training (I1), understanding of expert performance (I2), generalization to future situations (I3), and assessment and display of competence within the social setting of a debrief (S1).

Performance Diagnosis

The effectiveness of a debrief hinges on an instructor's ability to support trainees in diagnosing the underlying causes of mission failure (or success), and attributing those causes directly to individual or team behavior (D1). As noted above, serial mission replay capabilities are usually provided as a global representation of mission performance, which trainees observe to recall and learn the general flow of the mission. Traditionally, such replays do not include views of specific displays or instruments, which can provide the needed context of individual constraints and reasoning in diagnosing mission performance. When the training environment is extended to include multiple, diverse, and distributed training groups, the limitations of this training approach become increasingly apparent.

Because distributed training exercises eliminate the logistics and cost issues associated with live training events, more diverse trainees (or more trainee groups) can participate. For example, a distributed Naval training event may include an E2-C Hawkeye aircraft, designed to provide surveillance coordination, a set of F/A-18 Hornets, to provide suppression of enemy air defenses and strike missions, as well as ground control and intelligence support. The diversity of the participants in the training exercise, represented in this example by the two air platforms and the supporting ground elements, generates complexity in the type of mission that is executed, the data that is generated from the mission, and the interdependencies between actions in mission performance. With current procedures and technologies, instructors and trainees must diagnose successes and failures, and the interdependencies of actions to support these outcomes, using the replay of a simple common operational picture. Because this technology cannot capture the subtle details of actions and communications between *remotely* located elements, the diagnosis of performance is likely to be hindered.

The AAR process can be structured to encourage instructors and trainees to collaboratively identify performance shortcomings and strengths, under which instructors are an essential guide to the diagnosis process. Under a distributed training process, however, trainees and instructors will likely not all be collocated, thus inhibiting the element instructors from observing and integrating performance information in real-time across the package. In these cases, measures of performance may be collected automatically through the training simulation or remotely by observing performance through global views of mission performance. By distributing trainees and instructors across locations, instructors will also face challenges in diagnosing performance and ultimately may be less helpful in supporting performance diagnosis in trainees.

In order to support these diagnostic processes and leverage that diagnosis to improve future performance, it is critical that trainees and instructors are successful in recalling mission performance (I1), comparing that performance to expert behavior (I2), and extrapolating behaviors to future situations (I3). Each of these key processes can be inhibited by the dispersion of trainees across geographical locations, as we note below.

Performance Recall, Comparison, and Extrapolation

Traditional replay techniques can have great utility in supporting memory for the sequence of mission events (Schank, 1982) in traditional training events. Distributed training exercises can compromise this recall, by increasing the load placed on instructor and trainee memory, through two independent mechanisms. First, because distributed training exercises can involve richer interactions with heterogeneous training groups, the data and behaviors associated with these interactions increase in number and complexity. For example, communications in non-distributed training events may be limited to radio communications from a strike pilot to ground control. In distributed events, these same communications will occur, but may also be augmented with calls to pilots in supporting platforms. Second, distributed training exercises rely upon simulation systems that have evolved into sophisticated data collection tools, which can exponentially increase the amount of data that is collected (Jacobs, Cornelisse, & Schavemaker-Piva, 2006), and possibly recalled, by instructors. While all of these data may not be conventionally discussed in traditional debriefing processes, they may be included in distributed debriefs, thus increasing the demand on trainee memory and reducing the likelihood of recalling these granular data.

The comparison of these data, particularly for performance failures, to expert alternatives is further complicated by the complexity of the interactions between trainees. As the number of individuals involved in the training exercise increases, the predictability of their interactions and the ability to optimize or simulate these interactions in visual animations also becomes complex. Finally, the grouping of similar performance data becomes an essential component to understanding performance trends and extrapolating those trends to assessing future behavior. Because the distributed training environment provides more frequent opportunities for data collection, this training approach can reveal rich, informative patterns in trainee behavior. As in each of the critical AAR processes discussed thus far, the sheer quantity of data and the heterogeneity of behavior categories can complicate this process beyond that encountered with traditional training events.

Assessment and Display of Competence

The final function that the AAR provides is to support the social processes through which trainees can evaluate the competence of their colleagues, as well assert their own expertise. Whereas distributed debriefing processes can influence the AAR functions described above by generating more complex and diverse performance data, the distribution of trainees and instructors during debrief constrains the social mechanism of appraisal quite differently.

Distributed debriefing, and the tools that are used to support communication during these processes, can affect social appraisal and display of competence in three key ways. First, information sharing tools (e.g., collaborative desktops) and communication channels (e.g., video teleconference) can be useful in exchanging knowledge across disparate locations; however, these tools can still constrain the type of information that is readily shared. These tools are especially useful at exchanging text-based, or pre-generated spatial content, but are not highly effective at facilitating the interaction between trainees who are collocated at a whiteboard and sketch out mission developments and aircraft formations. Second, the microphones and sound quality associated with teleconferences or video teleconferences can limit the ease with which discussion can readily take place, even with the most advanced systems. These communication technologies are also poor at transmitting radio communications that are replayed during a mission review. Finally, these communication devices cannot capture the gestures and nonverbal communications that can be used to effectively assess and display competence. While traditional training processes can rely on nonverbal information exchange, such as eye contact, distributed training environments strip this form of communication from the essential social interactions that occur between individual trainees, as well as between trainees and their instructors.

Requirements for Distributed AARs

In the preceding section, we identified several challenges that distributed debriefings might, and frequently do, encounter. However, standard procedures for conducting effective distributed debriefing with collaborative technologies have not yet been defined. While a range of processes could be used to prepare and deliver debriefs in the distributed environment, the utility of these approaches depends both on the context in which they are being used, and the design of the training technology to support the instructor. It is tempting to discuss the technical and procedural requirements for distributed AARs separately. However, due to their very nature, distributed AARs combine both technology and process in a way that is difficult, if not impossible to separate. To ensure that the technologies and the processes interact seamlessly, it is important to address the design of these pieces simultaneously and collaboratively. We discuss distributed AAR requirements below, relating them to the previously defined AAR functions ((D1) diagnosis of performance, (I1) recall of performance training, (I2) understanding of expert performance, (I3) generalization to future situations, and (S1) assessment and display of competence.) Importantly, while these requirements may be met by a single technology, it is perhaps more likely that a series of AAR tools and institutional processes will ultimately fulfill distributed AAR needs.

Communication

A distributed AAR must support communication between the various locations involved, during both the preparation and delivery phases. At a minimum, voice communication should be supported. Ideally, video conferencing will be supported across all sites. This is no small technological feat, particularly across multiple locations. While current technology exists that supports this requirement, there are numerous technical issues that almost always pose some issues. Naturally, the frequency of technical issues increases with more distributed sites. Aside from technical issues, each distributed exercise must decide on and follow a few ground rules for using these communication technologies, beginning with the start of the communication (i.e., who calls who) and including turn taking, tips on reducing extraneous noise, and how electronic information will be shared. A critical requirement of any distributed AAR, all of the previously defined AAR functions are supported through this.

Collaboration

Closely tied to communication is collaboration. The distributed AAR must allow sites to coordinate on the content that should be debriefed, the strategy for debriefing, and the actual delivery. Thus, any performance data, simulator replay feeds, or other supporting information available must be shared across sites. Collaboration technologies that allow all participants to view the same information reduce confusion and facilitate the creation of common ground between all instructors and trainees. Optimally, the collaboration technology will allow each site to take control of the information, manipulating it ways that help illuminate the results of the training exercise. Here again, ground rules on effective use of the collaboration technology are imperative. Confusion in turn-taking can quickly lead to a chaotic AAR, both in the preparation and delivery phases. As with communication, without collaboration mechanisms, none of the AAR functions can be fulfilled in a distributed fashion.

Automated Data Capture

Distributed exercises run on a very tight schedule. Instructors are not given much time to develop an AAR; 20 – 30 minutes is fairly standard. Any distributed AAR technology must facilitate rapid AAR development, part of which is accessing available performance data and simulator feeds. Allowing instructors to view performance data specific to their element and common across all exercise participants supports diagnosis, recall, understanding, generalization, and, ultimately, overall performance assessment. The distributed AAR technology should accept and process performance data and simulator feeds automatically or semi-automatically (in the case of observer measures), with little direction on the part of the instructor.

Data Presentation

Any performance data (i.e., performance measurements) collected during the exercise have two potential presentation audiences: the instructors (during AAR preparation) and the students (during AAR delivery). How this performance data is presented can greatly facilitate or hinder the instructors' and students' ability to diagnose, understand and, subsequently, assess their performance during the exercise. The distributed AAR technology should allow the performance data to be presented at varying levels of detail, as it relates to the community, the package, and each element. Drill down capabilities are key, as are the methods by which performance data are presented (e.g., on a timeline, by event, textual representations, graphs, etc.).

Data Selection

Not all performance data is relevant for every element. Not all performance data is relevant for the entire group of exercise participants. In order to facilitate diagnosis and understanding of performance, instructors must be able to easily select the relevant performance data and simulator feeds or displays. Thus, the instructors must be allowed to identify key aspects of the exercise that are indicative of both good and poor performance. Subsequent review of performance during these key events can facilitate debriefs at the element, package, and community levels.

Replay Perspective

While viewing performance data is important, replaying exercise events is equally critical. Showing these events from multiple points of view can greatly facilitate diagnosis, recall, understanding, generalization, and assessment of trainee performance. Viewing events on a tactical map (or from a god's eye view), from a first person view, or a third person view can provide context and perspective to participants that would, otherwise, be difficult to obtain. Similarly, it may be useful to display gauges and instrument panels in order to provide a more common understanding of element capabilities across trainees. The distributed AAR technology must allow instructors to replay selected exercise events from these multiple points of view, to the extent they are available. Additionally, the distributed AAR technology must allow instructors to choose when and how these view points are presented, in order to best fit in with the overall AAR.

Expert Models of Performance

Viewing and analyzing performance in comparison to defined standards or models of expert performance can greatly assist instructors in diagnosis trainee performance. Discussing how the trainees' performance compared to these standards can assist the trainee understand what went well or poorly and understand why. The distributed AAR technology should present alternative or expert models of performance for each platform, either in the form of quantitatively modeled behavior or in the form of performance categories.

Flexible Delivery Style

The way in which an AAR is conducted varies according to the institution sponsoring the training exercise, the domain(s) being trained, and the instructors conducting the training. It is therefore important that any distributed AAR technology not unduly constrain or force instructors into presenting feedback to trainees in a specific style. This requirement can be tricky to fulfill. Certainly, some instructional strategies may be more effective than others. And indeed, some instructors may be more effective than others. However, all other things being equal, instructors should be allowed to tell the performance story in the manner that best suits their needs. For example, it may be appropriate to conduct the AAR in a way that tells a story, while allowing the instructor to drill down to specific training objectives and aspects of performance when appropriate. In this way, the AAR can be tailored to each trainee audience, thereby maximally supporting trainee learning.

Post-Exercise Review

Once the distributed training exercise is complete, the AAR should be available for off-line review by each site. Additionally, each participant should receive an AAR take-away report after the exercise is complete. This can facilitate a variety of post-exercise analysis activities, such as a more in-depth review of that site's performance, evaluation of training effectiveness, evaluation of instructor effectiveness, evaluation of AAR effectiveness.

Store Lessons Learned

Distributed exercises accomplish much more than teaching individual trainees. Each exercise results in a variety of lessons learned (e.g., about trainees, about training materials and scenarios, etc.) to be used when planning the next distributed exercise, in local training by units and squadrons, and during real-world missions when deployed. Retaining and using institutional knowledge is a difficult process in any environment. The distributed AAR technology must facilitate institutional learning by providing a mechanism to accumulate these lessons learned, distributed them, and use them when developing the performance assessment plan for the next exercise.

Scaleable

Distributed simulation training exercises come in all shapes and sizes. At their simplest, they involve two different elements, each located at a different geographic location. The elements may or may not be similar. At the other end of the spectrum, large scale distributed simulation exercises, such as the Air Force's Virtual Flag, may involve up to 25 different sites, involving activities at the operational and tactical levels. Adding to the complexity, some sites may host a variety of different types of simulators. Distributed AAR technology must be scaleable to many locations of the same or different platforms. It must work equally well with two sites (as in a package debrief) as with 25 sites (as in a mass or community debrief). Given the limitations in resources, it is simply not feasible for sites to learn and maintain multiple technologies and processes for conducting distributed AARs.

Ease of Use

Instructors may be involved in a distributed exercise a couple of times a year – if they're lucky. Time spent immediately prior to the exercise typically focuses on reviewing the relevant techniques, tactics, and procedures, becoming familiar with the scenario being executed during the exercise, and ironing out any kinks in the distributed technology. There is typically little time to become familiar with *how* the distributed technology works. Thereforem, any distributed AAR technology must be very easy to use, allowing the instructors to quickly learn (or re-learn) the technology with minimal hassles or instruction.

Current Techniques for Debriefing Distributed Teams

Many existing simulation environments provide the technical building blocks upon which many of the requirements listed above can be fulfilled. While some of these requirements have been fulfilled by existing technologies, no fully implemented distributed AAR system addresses all of these requirements.

State of the Art in Distributed Debriefing

Today's distributed simulation training has come a long way. Both the Navy and the Air Force regularly conduct complex virtual training exercises that involve a variety of simulated platforms and locations (e.g., the Navy's Operation Brimstone and the Air Forces' Virtual Flag Exercises). Research and development efforts across all of the U.S.

military services are continuing to develop additional simulation environments for use locally and in a distributed manner.

Large-scale Distributed Simulation Training Exercises

During large scale virtual training exercises hosted by the Air Force and Navy, the distributed AARs heavily rely on common tools found throughout the services. Standard video teleconferencing (VTC) applications are used to connect sites during briefs and debriefs. What formal data collection occurs is generally presented to all trainees in slide show form. During the debrief, collaboration technology (such as Information Workspace) is used to share these slide shows across the network and saved to commonly assessable shared network drives. The diagnosis of performance and delivery of the AARs is typically left up to the individual instructors, although some overall guidance on areas of good and poor performance may be provided from an internal assessment team. The slide shows largely contain textual information, with some still images as applicable.

This method of developing and conducting distributed exercises certainly has its benefits. The use of commonly available applications minimizes the maintenance required by each site and minimizes the need to learn yet another application by already busy instructors. Additionally, the free-form nature of the slide show allows the instructors to add whatever content is desired. On the other hand, these commonly available tools do not allow instructors to take advantage of the plethora of simulation data available to them. To do so would involve much more time than instructors currently have. A technology more focused on providing instructors with immediate data and feedback regarding the trainees' performance may help them develop the AAR more rapidly, covering more instructional points than they are currently able to.

Small-scale Distributed Simulation Training Exercises

Smaller-scale, distributed AARs are experimenting with distributed AAR technologies a bit more. Distributed AAR technologies such as the Navy's DDSBE AAR (developed by Aptima, Inc.) and AFRL's Distributed Mission Training Collaborative Briefing and Debriefing System are two such examples, primarily used to conduct a distributed AAR across two sites (Salter, et al., 2005, Sidor, 2002). Both systems support communication and collaboration across sites, replay of events using a tactical map, interaction with that map, and view of various instrument gauges. In addition, the DDSBE AAR (shown earlier in Figure 3) collects and presents formal performance data collected during the training exercise. Instructors are able to select specific performance data points for discussion and presentation during the AAR. To facilitate this task, the performance data are color-categorized into good performance (green), poor performance (red), and average performance (yellow).

These smaller scale distributed AAR tools are certainly on the right track for becoming a scaleable solution that can be used during large-scale distributed training exercises, as well. Additional work needs to be done to ensure that collaboration methods are scaleable across multiple sites, to incorporate multiple view points in replay capability, to ensure that formal performance data are collected and available for all trainees, and to support use of the data after the exercise is over.

Summary

Currently, most developers of distributed AARs focus either on the technology or on researching academic issues surrounding debriefing processes. Little information is publicly available describing efforts to combine these two strands in meaningful ways. In order to be truly successful, distributed training exercises must conduct AARs that rigorously and formally promote trainee learning. We believe that, by considering the current challenges to these issues and the requirements presented above, distributed AAR technologies and techniques can begin to meet this lofty goal.

References

Dismukes, R. K., McDonnell, L. K., & Jobe, K. K. (2000) Facilitating LOFT debriefings: instruc-tor techniques and crew participation. *International Journal of Aviation Psychology*, 10: 35.

Dixon, Nancy M. (2000). *Common Knowledge: How Companies Thrive by Sharing What They Know*. Cambridge: Harvard University Press.

Dwyer, D. J., Fowlkes, J., Oser, R. L., & Salas, E. (1996). Case study results using distributed interactive simulation for close air support training. In *Proceedings of the 7th International Training Equipment Conference* (pp. 371-380). Arlington, VA: ITEC Ltd.

Jacobs, L., Cornelisse, E., & Schavemaker-Piva, O. (2006). Innovative debrief solutions for mission training and simulation: Making fighter pilots training more effective. In *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*. Orlando, FL.

Salter, W.J., Hoch, S., and Freeman, J. (2005). Human Factors Challenges in After-Action Reviews in Distributed Simulation-Based Training. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting.*

Schank, Roger C. Dynamic memory: *A theory of reminding and learning in computers and people*. Cambridge University Press, 1982.

Shebilske, W., Gildea, K., Freeman, J., and Levchuk, G. (In press). Optimizing Instructional Strategies: A Benchmarked Experiential System for Training (BEST). Theoretical Issues in Ergonomics.

Sidor, G. J., Capt. (2002). Distributed Mission Training Collaborative Briefing and Debriefing System [Electronic Version]. *AFRL Horizons*. Retrieved August 6, 2007.

Tulving, E. (1984). Precis of Elements of Episodic Memory. *Behavioural and Brain Sciences*, 7, 223 – 268.

Tulving, E., & Schacter, D.L. (1990). Primary and Human memory systems. *Science*, 247, 301 – 306.