THE USERS AND FUNCTIONS OF DEBRIEFING IN DISTRIBUTED, SIMULATION-BASED TEAM TRAINING

Jared Freeman, Ph.D., William J. Salter, Ph.D., and Susannah Hoch
Aptima, Inc.
Woburn, MA & Washington, DC

Debriefing is the principal forum in which teams receive feedback concerning their performance in scenario-based training. We consider the audiences and functions served by current debriefing practices. We define additional audiences and functions that will be supported by an advanced debriefing system now being designed for Navy training exercises.

INTRODUCTION

Expertise is principally a function of amount of deliberate practice, that is, hours of well-structured, attentive exercise on relevant tasks. This was the finding reported a decade ago by cognitive psychologists Ericsson & Charness (1994) on the basis of research in several domains. It is an elaborate version of the punchline to the much older query, uttered by a musician lost on the streets of New York, "How do I get to Carnegie Hall?" (The answer, of course, is "Practice.")

The physical and behavioral realism of simulations facilitates engaging practice, but not necessarily well-structured, or relevant practice. Two techniques are typically used to enhance the value of practice during simulation. The first is principled design of scenarios to address mission essential competencies. The second – and the focus of this article – is debriefing (or After Action Review) in which trainees focus on important practice events and deliberate about them.

The feedback delivered in debriefs is particularly important in training teams and organizations (teams of teams) because teamwork itself often does not reliably produce immediate feedback (during task execution) from which team members can learn. Teamwork is typically distributed over time and space, and designed so that one team member can trap and repair the errors of another. As a result, the success or failure of one team member can be difficult for a distant teammate or trainer to observe immediately (if at all), and difficult to associate with actions made many minutes earlier. One faulty teamwork action may produce no observable failure at all if it is corrected by a teammate or rendered irrelevant by unexpected events. Debriefing can focus trainee attention on these fleeting teamwork behaviors. Like good scenario design, effective debriefing can substantially enhance the pedagogical impact of simulation-based practice.

But how can we help ensure that debriefing is effective? This is the question addressed in a science and technology program directed by the Office of Naval Research and NAVAIR Orlando, and conducted by its contractors: Aptima, CHI Systems, Lockheed Martin, MicroAnalysis & Design, and Sonalysts. In this article, we take initial steps towards defining effective debriefing. We first examine two studies of aircrew debriefing practices. We then define the audiences that debriefing systems must serve and the functions it must provide to them. We illustrate several core functions as they are implemented in a prototype debriefing system now in development.

CURRENT PRACTICE IN DEBRIEFING

Dismukes, McDonnell, and Jobe (2000) observed 36 debriefings of 2-3 person aircrews from five commercial airlines engaged in Crew Resource Management training. These industry-standard debriefing sessions began with a review of the agenda and expectations of participants, then continued with replay and instructor-facilitated discussion of audio and videotapes of the scenario run. Sessions ran 30 minutes on average, with 10 minutes devoted to scenario replay. The authors also administered a survey instrument to debriefing participants.

Analysis of data from observations and the instrument indicated that debriefs often failed in several ways. Many instructors did not engage trainees in setting the debriefing agenda or explain to them the need to candidly and actively critique their own performance. Instructors asked ineffective questions (questions that did not elicit critical or insightful comments from aircrew). Some abandoned facilitation in favor of lecturing to aircrew; they asked little and talked a lot. In addition, the authors found that debriefing was effectively delayed by hours because teams typically conducted several simulator sessions before debriefing the first. This attenuated instructor and trainee recall for the first run.

Jennifer Fowlkes has described large-scale military debriefing practices at the NSAWC Air Wing Training Detachment (Freeman, Littleton, Haimson, Stretton, Fowlkes, Carolan, Bilazarian, Radtke, 2003). These debriefings have three phases. In the preparation phase, instructors assemble to replay the scenario, diagnose performance, develop talking points, select illustrative sce-

nario video, and identify those diagnoses so uncertain that they require trainee input.

In the mass debriefing phase, the training leader (known as the "Overall") reviews the training objectives and mission plan, and then states the "Debriefing Rules of Engagement" (e.g., anyone may stop the taped replay to comment). The mission leader (or strike lead) critiques the planning and briefing process, after which instructors do the same. Mission video and audiotapes are then replayed as the mission leader facilitates discussion. The Overall periodically stops replay (at times selected during debrief preparation) to query the mission lead or others, reveal enemy views of mission entities and events, or comment (with other instructors) about team performance. Finally, the Overall assesses mission effectiveness, the mission lead assesses general performance of each team, and VIPs in the audience offer comments.

In the third phase of debriefing, mission elements (small groups of aircraft or other assets with unique mission roles) review their own performance. The format of these debriefings is specific to the mission element and often informal.

Users & Functions of Debriefing Systems

The reports described above reveal several important users and functions of debriefings. Trainee teams generate and receive performance feedback. Trainers review and diagnose mission performance, assemble debriefs, and administer them. Mission leaders assess mission outcomes and run debriefs. This list of users and functions is essential but not exhaustive. Here, we extend this list and illustrate its influence on a prototype system for Debriefing Distributed Simulation-Based Exercises (DDSBE). Most of the functions described below are supported by an advanced data collection system implemented in the current DDSBE system. Its components are described in other papers in this session.

Debriefing systems must help teams review feedback and generate it. This is the primary function of debriefing systems. It is complex because debriefs must support trainee-generated discovery learning during replay, trainers' didactic commentary during replay, and trainers' analysis by factors other than time, including training objectives, mission objective, and trainee groups.

Figure 1 shows one display of a prototype debriefing system that supports these functions. The display shown is designed to support trainers who assemble debriefs. A similar display (less the check boxes to the right) will serve trainees.

On this display, mission replays will present mission events unrolling over time using an animated tactical map, student communications audio replays, transcripts of expected communications (on the lower left in Figure 1), and, potentially, dynamic instrument readings. Mission replays like this are the dominant form for debriefing because it enables trainees (supported by their leaders and trainers) to review scenarios in the narrative order in which they unfolded and in which they may be organized in memory (Schank and Abelson, 1977). This is also the mission order in which practiced skills are likely to be applied in real life, so studying scenario events in their simulated sequence may prepare trainees to experience and execute them in their natural sequence.

In support of didactic commentary by trainers during replay, the system flags moments of teamwork failure (e.g., omitted or incorrect communications, errors in transmitting data or orders between warfighters). A red warning in the lower left pane marks one such failure.

Finally, the system supports feedback structured by factors other than time. The upper right pane of Figure 1 shows a hierachical index into the training objectives of the scenario. Each objective is automatically scored (using small green and red graphs). As the trainer selects an objective, the replay is automatically positioned at a corresponding moment in the scenario. Thus, abstract scores on training objectives (or other factors) are explained by specific evidence from the mission replay. Similar hierarchies exist for the mission events and mission teams. Users can select whatever view is most useful to understand performance.

Debriefing should teach trainers to facilitate trainee critiques of their own performance. Teams that think critically about their own performance learn better (Chi, et al., 1994; Cohen, Freeman, & Thompson, 1998). The Army has an admirable tradition of encouraging this behavior in its After Action Reviews, and sound guidelines for conducting them (Dixon, 2000): "No sugar coating, discover ground truth, no thin skins, take notes, and call it like you see it."

In future development of DDSBE, we will explore ways of presenting instructors and leaders with questions with which to prompt trainees to critique the team's performance and their own. This should encourage participation by team members, reveal useful observations not observed by trainers or measured by the system, allow future leaders within the team to show themselves, and help leaders and trainers learn to facilitate discussions rather than dominate them.

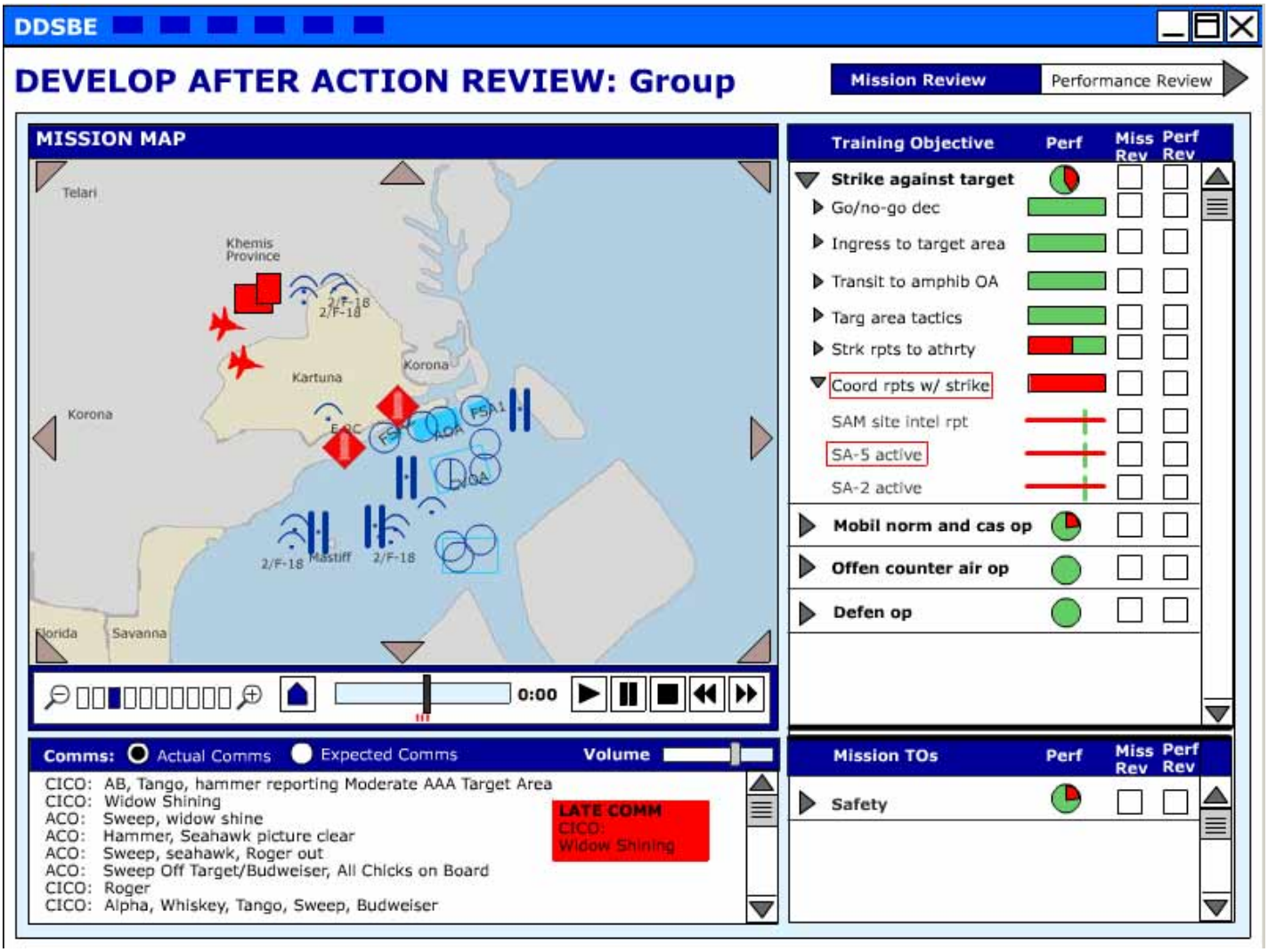

Figure 1: After Action Review. The left hand portion of the debriefing display shows an animated mission replay. The right hand portion presents performance on various training objectives and an index into the replay for each objective.

Debriefing systems must identify and support emergent teams. Debriefing can serve two types teams. Currently debriefing practices serve pre-defined teams, such as mission elements, who join together during debriefing to analyze their internal coordination, their coordination with outside elements, and to build team morale and team-awareness.

However, technology-supported debriefs can also serve emergent teams, combinations of warfighters from different elements and platforms whose interactions in the particular scenario hampered the mission. It is currently possible, but difficult, for trainers to identify such teams manually during scenario play, and almost impossible to assemble them (even virtually) on short notice afterwards, for debriefing. In later phases of this project, we hope to extend organizational modeling techniques previously developed to design command and control teams (MacMillan, Paley, Levchuk, Entin, Serfaty, & Freeman, 2002) to automatically identify emergent teams. These teams can then assemble using collaboration technologies to conduct debriefs of their most critical interactions (see Hall-Johnston, Serfaty, & Freeman, 2003).

Debriefing systems must help distributed teams collaborate. As suggested above, the distribution of teams poses significant challenges for all of the participants in debriefing. Training staff (including trainers, observers, and controllers) must collaborate to analyze scenario events from their different perspectives and come to agreement concerning the feedback and evidence to present to trainees. Similarly, trainees must pool their perspectives on the training mission. All of this must be done across great distances – between ships, schoolhouses, simulation control rooms, and other facilities.

Our initial design to support this collaboration features commercial off-the-shelf applications for audio chat, text chat, whiteboards, and application sharing.

Whether these tools are useful and used are matters for study over the duration of this program.

Debriefing systems should help instructors to assemble and deliver training debriefs. The blessing of simulation-based training – abundant data – is also its curse. The DDSBE system is instrumented to capture large volumes of performance data from observers (using a handheld device) and trainee workstations. Even after aggregating these data and assessing them, there is more information available for debriefing than there is time to present it. Nor is there reason to discuss much of it. It is the exceptional failures (e.g., collateral damage) and successes (e.g., evading enemy air defense) that typically concern instructors and trainees. Checkboxes to the right of Figure 1 are one simple device with which instructors will select the content (e.g., the scores on training objectives and associated scenario events) to deliver to trainees in debriefing.

In addition, the quality and strength of diagnoses that instructors develop while preparing debriefs should improve using this system. Abstract scores (on the right of the display) and the scenario narrative (on the left) index one another in a manner that should enable instructors to diagnose failures more quickly and accurately, and to quickly compile evidence that strongly supports and illustrates those diagnoses.

Debriefing systems should help mission leaders learn to conduct mission debriefs. Mission objectives, like training objectives, are represented in the hierarchy on the right of the prototype display, and they are also indexed to the mission replay on the left. This functionality is intended to help leaders focus team attention on mission goals, while providing evidence of success or failure through mission replay.

Debriefing systems should help instructional designers and researchers to evaluate and improve training. Debriefing is an opportunity to manipulate and analyze the materials, structure, and process of delayed feedback in a quasi-experimental manner. Findings from such studies may inform debriefing in training as well as mission debriefs in military operations.

In future research in the DDSBE program, we hope to develop additional controls for selecting, ordering, and authoring debriefing content, and methods of measuring the impact of those manipulations on learning.

In addition, we plan to instrument this debriefing prototype to gather data on the use of its features, and the perceived validity and value of scenario events, performance measures, and feedback. These data can be gathered by logging which scenario events, measures, and feedback leaders, trainers, and trainees attend to in debriefs. In addition, we hope to elicit users' subjective feedback concerning the validity of debriefing components (especially assessments) and their perceived value. These data can then guide improvements to scenarios, measures, and feedback.

CONCLUSION

We have identified several users of debriefing systems – not all of them obvious: trainees, trainers, operational leaders, and instructional designers/researchers. We have defined a number of functions of debriefing systems that can serve these users. The prototype debriefing system for DDSBE illustrates ways in which some of these functions can be realized.

As we develop this system, our principal goal is to create a debriefing system that improves training for distributed teams, one that delivers valid assessments generated by DDSBE and culled by trainers, and one that provides trainees with opportunities for deliberate review of team performance in multiple ways.

ACKNOWLEDGEMENTS

This work is supported with funding from the Office of Naval Research. The ideas expressed herein are those of the authors, and are not necessarily endorsed by the funding agency. This material is approved for public release; distribution is unlimited.

REFERENCES

Chi, M. T. H., de Leeuw, N., Chiu, M. H., & La-Vancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, 18, 439-477.

Cohen, M.S., Freeman, J.T., & Thompson, B.T. (1998). Critical Thinking Skills in Tactical Decision Making: A Model and A Training Method. (Canon-Bowers, J. and E. Salas, eds.), *Decision-Making Under Stress: Implications for Training & Simulation*. Washington, DC: American Psychological Association Publications.

Dismukes, R. K., McDonnell, L. K., & Jobe, K. K. (2000) Facilitating LOFT debriefings: instructor 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.

Ericsson, K. Anders & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist*, 49(8), 725-47.

Freeman, J., Littleton, B., Haimson, C., Stretton, M., Fowlkes, J., Carolan, T., Bilazarian, P., & Radtke, P. (2003). *Concept of operations & experimental plan (v2.0) for debriefing distributed scenario-based exercises (DDSBE).* Ap-

tima Technical Report AP- 1219. ONR Contract #N61339-99-D-0012.

Johnston, J.H., Serfaty, D., & Freeman, J. (2003). Performance measurement for diagnosing and debriefing distributed command and control teams. *Proceedings of the 8th International Command and Control Research and Technology Symposium*. Washington, DC.

MacMillan, J., Paley, M.J., Levchuk, Y.N., Entin, E.E., Serfaty, D., & Freeman, J.T. (2002). Designing the best team for the task: Optimal organizational structures for military missions. In M. McNeese, E. Salas, & M. Endsley (Eds.), *New trends in cooperative activities: system dynamics in complex settings*. San Diego, CA: Human Factors and Ergonomics Society Press.

Schank, R.C. & Abelson, R. *Scripts, Plans, Goals, and Understanding*. Lawrence Erlbaum Associates, Hillsdale, 1977.