Training Novices and Experts: A Common Assessment Mechanism for Knowledge, Skills, and Abilities

Mary Ann Pigora, Theresa Tamash
Applied Research Associates, Inc.
Southeast Division, Orlando, FL
mpigora@ara.com, ttamash@ara.com

Holly C. Baxter, PhD.
Applied Research Associates, Inc
Klein Associates Division, Fairborn, OH
hbaxter@ara.com

ABSTRACT

There are a number of assessment data models that support authoring of objective (question-based) tests. Many of these include an intelligent tutor that can present “hints,” remedial material presentation, or advanced placement, providing tailored feedback in a timely and cost effective manner. Developing effective assessment and decision support models in the “free play” environment of a simulation-based exercise is more difficult, and disparate assessment models can result in inconsistent training. Current research in Naturalistic Decision Making (NDM) investigates the strategies people use in performing complex, ill-structured, and high-stakes tasks under time pressure, uncertainty, and in the context of organizational constraints. In many dynamic, uncertain, and fast-paced environments, there is no single right way to make decisions. Thus, the NDM approach typically studies experts to define quality decision making and describe good decision-making processes. This paper outlines an objective system to teach and assess these NDM skills, both for the individual and as a team.

The goal of our system is to train an individual from novice level to expert starting with little or no exposure to the target domain. We will define our system in three phases: knowledge, skills, and abilities, to correlate with Bloom’s taxonomy of learning domains. In the knowledge stage, the novice gains understanding of the domain. In the skills stage, the trainee translates knowledge into behavioral demonstrations of the material. In the abilities stage, the trainee applies the skills to make decisions in a real world team environment with uncertainty, time constraints, and organizational constraints. The system will assess all phases from a single knowledge base, providing consistent training across multiple presentation methods and levels of expertise and helping students build a pattern base that allows them to better make recognitional decisions when real world challenges arise.

ABOUT THE AUTHORS

Mary Ann Pigora is a Senior Scientist at Applied Research Associates, Inc., where she conducts emergency response simulation and training efforts. She has over 12 years of experience in 3D graphics and simulation, and is the lead architect on the Automated Exercise and Assessment System. Previous experience includes creating CGF and visual databases for human simulation, work on the OneSAF Objective System, and computer animation for the movies *Mulan* and *Fantasia 2000*. She received her BS in Compute Science from the University of Central Florida and her MS in Computer Science from the Georgia Institute of Technology.

Theresa R. Tamash is a Scientist at Applied Research Associates, Inc. For the past six years her work has been focused on simulation training systems within both the military and civilian domains, including Close Combat Tactical Trainer (CCTT), OneSAF Objective System, and the Automated Exercise and Assessment System (AEAS). Ms. Tamash obtained her MS in Computer Science from the University of Central Florida and her BS in Computer Science from Rollins College.

Holly C. Baxter, Ph.D., is a Senior Scientist at Klein Associates, a Division of Applied Research Associates, specializing in Instructional Design, Evaluation Metrics, Organizational Development, and Training. Her research interests include Knowledge Management and all phases of cognitively based training including cognitive needs analysis, design, development, implementation, and developing cognitive metrics for evaluating training initiatives ranging from classroom training to simulations. Dr. Baxter holds a BA in Communication from the University of Dayton, a MA in Organizational Communication and Training from Indiana University, and a Ph.D. from Indiana University in Organizational Communication and Management with a focus on Training, Instructional Design, and Organizational Development.

Training Novices and Experts: A Common Assessment Mechanism for Knowledge, Skills, and Abilities

Mary Ann Pigora, Theresa Tamash
Applied Research Associates, Inc.
Southeast Division, Orlando, FL
mpigora@ara.com, ttamash@ara.com

Holly C. Baxter, PhD.
Applied Research Associates, Inc
Klein Associates Division, Fairborn,
hbaxter@ara.com

INTRODUCTION

Browsing through brittle 1950s era textbooks, seeing dense columns of text broken only by simple line diagrams or poorly reproduced black and white photographs, it is clear that today's technology can offer a richer, more engaging learning experience. However, e-learning has not lived up to this ideal. This is not the fault of technology, but of its use in course design and execution. In a wake-up call to the e-learning industry, Frank L. Greenagel states "… the e-learning experience is, far too often, puerile, boring and of unknown or doubtful effectiveness." (Greenagel 2002)

In order to create effective e-learning systems, developers must understand how people learn (Greenagel 2002). Welford (1968) articulates steps for skill acquisition that include perceiving and comprehending material, holding the material in short-term memory and establishing a memory trace, then recognizing an appropriate situation in which to use the information. This model is easily implemented in an e-learning system. Introductory material (doctrine) is presented (ideally in an interesting and engaging format), and at the end of each unit of material a short review is given, often with a multiple choice quiz. Situations for applying the material are presented with a "comprehensive" test at the end. This is fine for novice-level knowledge where there is always a right answer. Real life isn't always so simple.

E-learning for the expert is much more complicated. A survey of research in decision theory results in many proposed models, but they all agree that experts have developed cognitive and perceptual skills to improve their decision making (Elliot 2005). Current research in Naturalistic Decision Making (NDM) rejects the classical analytical decision making models where an expert weighs each option and chooses the best one. The NDM perspective investigates the strategies people use in performing complex, ill-structured, and high-stakes tasks under time pressure, uncertainty, and in the context of organizational constraints (Klein, Orasanu, Calderwood, & Zsambok, 1993; Miller, Zsambok, & Klein, 1997). In many dynamic, uncertain, and fast-paced environments, there is no single right way to make decisions. Thus, the NDM approach typically studies experts to define quality decision making and describe good decision-making processes. Researchers using the NDM framework have examined expert performance with a wide variety of professionals in hundreds of studies. By studying the cognitive aspects of expert performance in a broad array of domains, NDM researchers have been able to make recommendations on how to improve training and system support to facilitate performance of non-experts (Klein 1998). NDM looks at a series of decisions in context, rather than a single moment of choice (Elliot 2005). How can e-learning teach—and assess—these skills effectively?

BACKGROUND

Gary Klein's Recognition Primed Decision-making model (RPD) focuses on how an expert compares a situation to previous experiences and finds it familiar, and from that recognizes a typical action to take. It is unlikely a previous experience will exactly match the current situation, so the expert focuses on assessing a situation and finding familiar patterns. The expert then runs through a mental simulation of the solution, looking for holes and finding workarounds. The final solution is a workable, good solution, but *may not be the best possible solution* (Klein 1998).

The clue to how e-learning can teach this skill set is found in the very method RPD proposes as the expert's litmus test: mental simulation. The expert engages in building a mental model of the situation and mentally trying out the solution. Computer-based simulation can put the trainee into a cognitively authentic time pressured situation, with shifting conditions, unclear goals, and incomplete information, the very situation that NDM seeks to study. Simulations give both novices and experts the opportunity to further construct their mental models, especially when the pattern based they have to draw on is limited.

Experts often work in teams to handle complex situations, such as crisis management teams. While the team leader is ultimately responsible for the decisions made, the team guides those decisions. Therefore when training experts, it is important to consider the team context in which the expert is immersed (Elliot 2005). Salas, Sims, and Burke's (2004) "Big Five" core competencies of teamwork show that effective teams have leadership and clear direction while providing coaching to develop members' team competencies. They monitor each other's performance to catch mistakes, and use shared mental models to handle workload distribution problems. They can recognize deviations from the expected and are flexible enough to adapt quickly. The members are team oriented, and know and trust each others' intents and competencies.

An effective method for gaining team cohesion is to participate in exercises that allow practice of these competencies and provides metacognition (thinking about how the team thinks) to evaluate their proficiency in handling a problem. Again, computer-based simulation can provide a time and cost effective team exercise methodology that allows the team members to carry out their individual roles, strengthen their coordination, and learn how to avoid breakdowns in communications, planning, and execution (Figure 1). Work has been done to define how human observer/controllers can assess team performance during computer "game play" (Weil, Hussain, Diedrich, Ferguson, & MacMillan). We will discuss an automated assessment component to provide an unbiased measure of team competencies, complementing the work of human observer/controllers and allowing them to concentrate on the finer points of the exercise.

REQUIREMENTS FOR A SUCCESSFUL SYSTEM

Training System Structure

The goal of our system is to train an individual from novice level to expert starting with little or no exposure to the target domain. We will define our system in three phases: knowledge, skills, and abilities (Figure 2), to correlate with Bloom's taxonomy of learning domains (Bloom 1956). In the knowledge stage, the novice gains understanding of the domain. In the skills stage, the trainee translates knowledge into behavioral demonstrations of the material. In the abilities stage, the trainee applies the skills to make decisions in a real world team environment with uncertainty, time constraints, and organizational constraints. One key advantage of this type of training is that it helps students build a pattern base or surrogate expertise that allows them to better make recognitional or intuitive decisions when real world challenges arise. The system will assess all phases from a single knowledge base, providing consistent training across multiple presentation methods and levels of expertise.

Figure 1. Simulation-based team exercise event at the Weber County, UT Emergency Operations Center in 2004. The AEAS system was used to present a sarin release scenario for participants to mitigate.

In the "knowledge" phase of learning, the trainee is exposed to the accepted doctrine of the domain. This exposure will teach the trainee the basic concepts and shared vocabulary of the domain, and can be accomplished through classic presentation-based training, structured and asynchronous learning networks, and objective tests. These systems are typical enough that standard system evaluation metrics have been defined (Valenti, Cucchiarelli & Panti 2002), however, it is still up to the course developer to make the content and presentation interesting enough to hold the attention of the trainee to successful completion (Greenagel 2002).

Figure 2. Our objective system will train an individual in a complex domain, starting with domain knowledge, then translating knowledge to behavioral demonstrations or skills, and finally applying those skills in a naturalistic decision making environment to achieve abilities.

After basic familiarization, the trainee must learn his role in the decision making process. By learning the skills required in his own job, he takes a step toward becoming part of a team identity (Klein 1998). Expertise is acquired through quality practice (Elliott 2005) and in this "skills" phase of training, the trainee can start building his repertoire of previous experience through individual computer-based simulation. In this phase, team members are simulated with enough fidelity to provide injects and interactions with the trainee. The trainee may try different roles in the simulation in order to gain an understanding of the responsibilities of other team members.

Finally, in the "abilities" phase, the trainees come together as a team to execute their roles in a collective simulation-based exercise. During this time, they should learn something about each other's roles, abilities, responsibilities, and lines of communication. Once they have practiced these skills to become a cohesive team, they can concentrate on the decision challenges facing the team as a whole (Klein 1998).

Training System Capabilities

In order to use e-learning to effectively train and exercise experts, our system must be able to discriminate between successful and unsuccessful decisions. In a structured learning environment, success is easily defined as giving the correct answer to a direct question. When learning is unstructured, as in the free-play environment of simulation systems, success is more difficult to define. Many simulation systems provide no measure of success, but it is difficult to build up expertise when there is less chance for concrete feedback (Shanteau 1992). Domains studied in naturalistic decision making such as firefighting or emergency management have real and obvious outcomes such as extinguishing a building fire with minimal damage. In order to use simulation-based training to develop expert abilities in our trainees, our simulation must be of a high enough fidelity to accurately convey these outcomes, otherwise the advantage of feedback is lost.

Assessing only the outcome, however, does not tell the trainee if success was due to his timely decisions or in spite of them. An alternate measure of success might be to determine how well doctrine was followed, both by individual participants and the force as a whole. While NDM indicates that effective decision making does not necessarily follow doctrine, in order to function in a cohesive team, each member must know his responsibility as defined in the applicable doctrine and team expectations.

The ability to automatically assess cause and effect will be a key advantage to our system. Conducting a comprehensive After Action Review (AAR) of a sophisticated exercise requires that a team of analysts and observer/controllers not only observe the exercise, but also evaluate the resulting reams of data. Often the AAR does not take place until weeks after the exercise when the experience is stale in participant's minds. Automated assessment allows immediate feedback, is more objective with less human bias and error, and allows the trainee to learn anytime, anywhere. Automated assessment of even basic skills will allow human observer/controllers to concentrate on the finer points of the exercise, increasing the value of their input.

The most difficult level of assessment that must be achieved by our system is that of the cognitive function of the team. Klein and Thordsen (1990) identify five cognitive functions that can account for many team errors: attention, both giving attention to important cues and ignoring chaff; motivation, or clear intent;

sensori-motor functions, including coordination of team members and knowledge of their ability to respond; reasoning, such as deriving inferences from information sources; and metacognition, which includes a grasp of the team's ability to handle the problem and strategies to handle team limitations. Poor metacognition can result in micromanagement that wastes resources. Our objective system must be able to access the team mind in order to assess experts exercising together.

PREVIOUS WORK

As noted previously, there are many systems available that offer assessment data models for structured (question-based) assessment. Previous work on assessing participants in simulation-based exercises includes Synthetic Teammates for Realtime Anywhere Training and Assessment (STRATA) and the Automated Exercise and Assessment System (AEAS). We will briefly discuss these systems and how their capabilities meet certain of our objectives.

Synthetic Teammates for Realtime Anywhere Training and Assessment (STRATA)

STRATA provides on-demand training and exercising for Naval aviators in Close Air Support (CAS) missions. Two roles are available for a human trainee within the aircraft, lead pilot and wingman. Either role, as well as support roles, can also be played by a simulated participant. STRATA's purpose is to assess trainees on meeting mission objectives, following procedures properly, and working with their teammates (Bell, Johnston, Freeman, & Rody 2004).

The Automated Exercise and Assessment System (AEAS)

AEAS is a computer-based simulation system for training and assessing command level emergency response teams, both in the Incident Command Post and in the Emergency Operations Center (EOC) (Figure 3).

Figure 3. AEAS is a simulation-based emergency management exercise system that provides role-based training in a team environment. AEAS provides an automated assessment of how well team members accomplished specific tasks.

AEAS can be run by a team, or by an individual with simulated teammates, and provides a broad collection of disaster scenarios. AEAS uses a system of Tasks, Conditions and Standards (TCS) to describe participant expected actions for each type of scenario. The TCSs are used to assess player actions, provide reminders and tutored training prompts, and to simulate missing role players (Pigora, Barshatzky, Kerrigan & Murphy 2002).

Foundations for Moving Forward

In this section we will discuss how current systems address assessment of individuals and teams, and training engagement and effectiveness. These are lessons for reuse as we move forward with our objective system.

Individual Assessment

STRATA uses simulated participants to test a trainee's situational awareness and adherence to procedure. Synthetic participants have various levels of competency to allow a trainee to gain experience working with rookies and experts. A simulated wingman may provide incorrect information. The trainee must recall data from a pre-mission briefing and be aware of the current status of the mission (have good situational awareness) in order to correct the wingman successfully.

STRATA's purpose is on-demand training, so these simulated teammates provide a substitute for real-life team interaction. The assessment mechanism has the advantage of knowledge of all actions taken by simulated participants. Our objective system will additionally attempt to assess interactions and coordination between human teammates.

Team Assessment

AEAS has some of the team assessment capabilities of our objective system. AEAS is role based, allowing team members to learn their responsibilities, which is a step toward an integrated team identity (Klein 1998). Each role is assessed on taking the steps to complete cognitive tasks such as keeping control of the scene, communicating both within the team and with outside agencies, and utilizing resources (Figure 4). Responsibilities may overlap between team members; as long as one person executes the tasking, everyone will get credit.

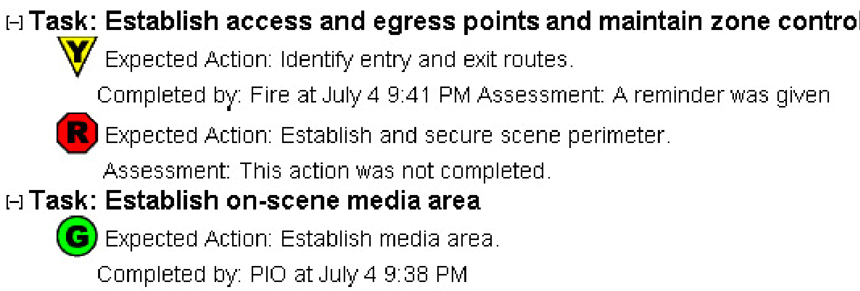

Figure 4. AEAS participants are given an assessment of Green, Yellow, or Red indicating they executed the subtask to standard, minimally to standard, or not to standard. In tutored training mode, decision support prompts for each subtask are displayed at the appropriate time to execute that subtask.

For the purposes of evaluating AEAS against our objective system, it is important to understand the fidelity of the task assessment. AEAS has no artificial intelligence. Subtasks that the participants are assessed on can be fulfilled by giving commands in the system such as sending a status report to another role player. When the report is sent, the assessment is checked off, but no attempt is made to determine if the report contents are useful. There are some specific exceptions such as notifying someone of a particular hazmat, in which the report form that is presented for the participant to fill out has checkboxes for the different types of hazards that can be reported.

While it is a valid beginning to know that the report has been sent or that a resource has been tasked with a relevant order, in order to assess the team mind we must understand whether the communication was effective or the tasking was in concert with actions taken by other participants. In determining the cognition of the team, AEAS falls short of our objective system.

Training Effectiveness

According to Greenagel (2002), corporations are more interested in throughput and low unit cost, so solid measures of effectiveness are infrequently developed or applied. No training effectiveness study has been performed for AEAS, however, empirical evidence indicates that the assessment does indicate the overall effectiveness of the team. The authors facilitated three exercises in one facility on successive days. On each day, thirty participants were first familiarized with the AEAS system, given a short training scenario with an explosion in a crowded venue, and finally participated in an assessed scenario with an explosion and sarin release at an outdoor concert in a park.

On the first day, the participants consisted of experienced emergency response professionals who were used to working together as a team. They handled the scenario calmly and "by the book." They received the best assessment of any group we had ever exercised, with virtually all green assessments. The second group was younger; many of them were the junior staff of the previous days' team. They didn't have the routine down as smooth as the previous group, but still received a good assessment. The third group was composed of trainers and emergency management officials from across the state, and did not know each other or work together regularly. They got so mired in the first twenty minutes of the scenario that we had to stop and let them discuss their command hierarchy before continuing. The assessment reflected their lack of coordination with mostly red assessments.

Training Engagement

One system that seeks to increase engagement of the trainee is the Interactive Storytelling Architecture for Training (ISAT). ISAT uses an intelligent agent, the director, to modify scenarios with injects or new scenes based on the trainee's performance (Magerko, Wray, Holt, & Stensrud 2005). The director's assessment of trainee performance is internal, and is not presented to the user directly.

Team interaction when solving a complex problem helps avoid the e-learning boredom zone and keeps participants interested. When there are still issues on the table, the team doesn't want to stop. The authors once called a grueling AEAS exercise after four hours to break for lunch, and the participants protested. As one participant wrote on the critique form for a three day training course, his expectation for the course was to be bored in front of a computer. In answer to whether his expectations were met, he wrote "Thankfully not!"

TECHNICAL DESIGN

Our objective system will be composed of five components: the assessment data model, which contains the assessment criteria data; the assessment engine, which examines user input and tracks performance against the assessment model; the assessment criteria authoring system, which allows a subject matter expert to turn doctrine into assessment criteria for the system; the presentation methodologies for the material or exercise; and the learning management system, which compiles AARs, tracks trainee performance and preferences over time, and presents the presentation methodologies (Figure 5).

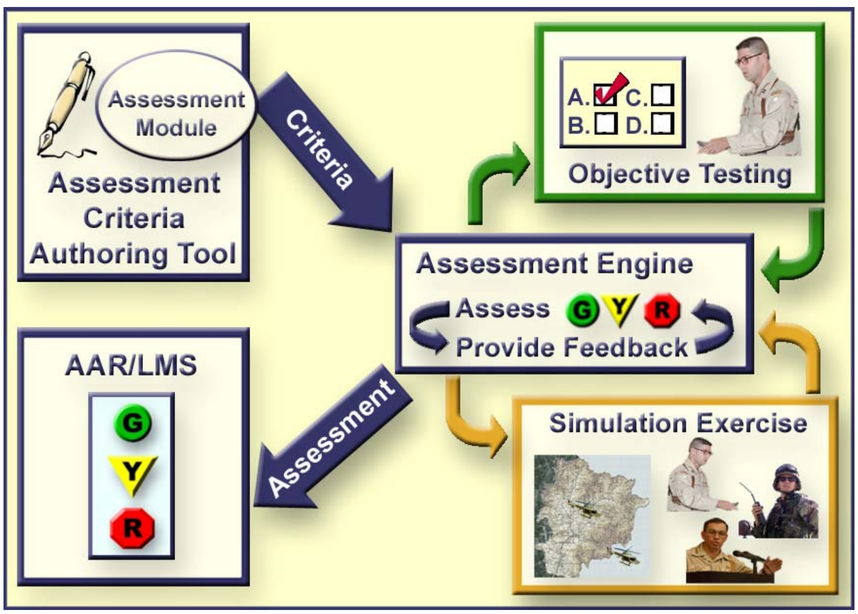

Figure 5. The objective system will consist of an assessment data model, which is instantiated as a domain assessment module using the criteria authoring tool. The module is used by the assessment engine to assess user input in different presentation methodologies.

Assessment Data Model

The assessment data model will be organized into individual skills. These building blocks will be grouped under cognitive tasks that must be accomplished during a scenario. An individual skill can be part of more than one task. Each skill will have a description, including one or more structured questions to test knowledge of the skill, events that can trigger the skill to be assessed in a scenario, actions or combinations of actions that fulfill assessment of the skill, criteria for successful application of the skill, such as a time limit or adverse events that indicate the time horizon has run out, and consequences in the scenario progression (either complications or remedial injects) if the skill is not correctly applied. The data model must also specify whether a skill is done only once in a scenario (such as setting up a command post) or is reset by particular events (such as sending a status report every day), and what role or roles are responsible for applying the skill.

Skills may have varying levels of fulfillment defined in the data model. A skill such as lowering aircraft landing gear can be assigned one of two assessments, completed or not completed, but a skill such as briefing subordinates may have varying levels of success associated with how well it is accomplished. The briefing could be given a different assessment depending on whether it conveyed all or part of the vital information, or reached all of the team members or a subset.

Assessment Engine

The assessment engine will apply the assessment data model to the trainee inputs. For a question-based assessment, this is trivial. In a simulation, however, we must find ways to track the cognitive processes using only the trainee's inputs into the system. For this reason, during an exercise the participants will be encouraged to use the communication channels they would normally employ in the real world, but we must introduce artificiality by requiring that all decisions be input into the system in the form of commands or reports. The form of the input will depend on the capabilities of the system interface and may include text entered with the keyboard, commands with buttons and checkboxes, and analysis of voice over IP for keywords. All participant input will be recorded for AAR purposes. In the skills phase of training, assessment is very similar to the model used in AEAS for individual skills. We will now propose methods for tracking the five cognitive functions of a team that are identified as sources for team error.

Attention

By defining our assessment criteria, we have defined what injects trigger the need for an action. These are the important cues, and the trainee(s) must respond to them appropriately. Conversely, injects that do not trigger the need for an action should not be taking up the team's time. By tracking the communications among the team the system can build a profile of which injects took the attention of the team. For example, we can compare the frequency with which trainees forward information from unimportant "chaff" emails with the communication of important issues (Figure 6). In the case of a very efficient team, attention to chaff may be acceptable as long as they have already taken care of the more important injects.

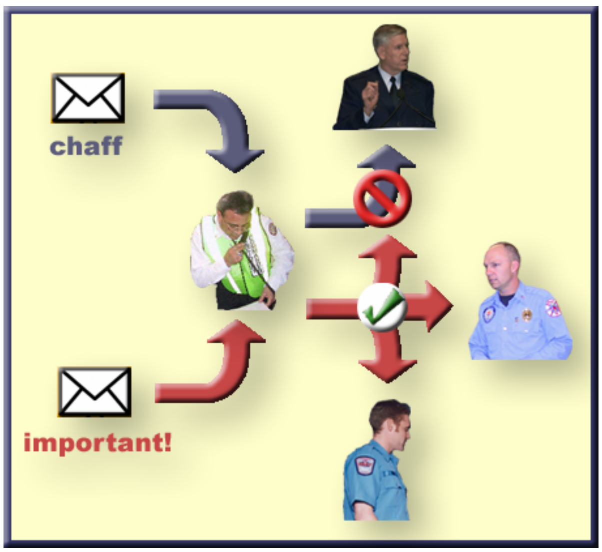

Figure 6. By profiling the communications between team members, our objective system can determine if the important injects are disseminated correctly and if the chaff messages are taking up valuable team resources.

Motivation

To assess motivation, the system must determine whether the team can define their goals and intents. In emergency management, overarching objectives are supposed to be set down in the Incident Action Plan, which defines the goals for a particular time slice of the incident. Being able to articulate concrete, achievable goals in a plan which is then disseminated to the team goes a long way toward team motivation. To aid in automated assessment, the team may also be prompted by the system at intervals to indicate what the leader's expectations are and what their plan is to achieve each goal. These commands would be compared to those needed to fulfill pending skills applications.

Sensori-motor functions

To determine the coordination of the team, the assessment engine can look for positive efforts to improve coordination. Good coordination examples might be information being volunteered proactively or requests for reports and information followed by a timely response. Once coordination is established, needed actions must be carried out. The team must understand how long it takes for actions to be executed. An indication of unrealistic perceptions of speed might be not giving commands in time to meet the skill time criteria or repeatedly asking a teammate if a task is complete before that task could realistically be accomplished.

Reasoning

Assessment of reasoning can be encoded into the scenario by creating particular circumstances that the trainees should infer from. An example is the common emergency response training ploy of using an initial event such as a small explosion to draw the responders in, followed by a larger attack to disable as many responders as possible. By failing to consider the possibility of a secondary attack, the team has performed an inadequate mental simulation of the response. In the example, expected skill application should include protecting the responders by deploying the bomb squad to search for devices and requiring the responders to wear protective gear in the incident area. Another way to assess reasoning is to provide trainees with conflicting information. Trainees must recognize that at least one of their sources is in error, and either ask for clarification or find the information through a trusted channel.

Metacognition

To operate effectively, the team must monitor its own performance, capitalizing on strengths and compensating for weaknesses (Klein 1998). By specifying what roles are responsible for certain skills and comparing this to the roles actually completing the tasks, our assessment engine can narrow inconsistencies to either stronger team members compensating for weaker, unequal task loading, or confusion among the team responsibilities. Highlighting inconsistencies in the AAR allows the team to hash out the roles and responsibilities.

Some insights about metacognition can also be gained by tracking the flow of information through the team, as discussed in the section on attention. Is there a very high message volume with a lot of repetition, indicating that information is being sprayed over the entire team and distracting members from their individual objectives? Is vital information not making it past the first person to receive it? Is the team noticing and responding to impending problems, i.e. when a task is triggered, does the team respond by applying skills? Is the team completing the basic procedures (individual skills) but in an uncoordinated way that results in high costs in the scenario outcome (excessive damage or loss of life)? In assessing the last item, a baseline scenario outcome, created by executing the team roles according to doctrine, can be used for comparison.

To achieve the highest level of automated assessment, our assessment engine must track user input with a high degree of granularity, which introduces some artificiality into the exercise. In an intense exercise, participants will talk face to face or on the phone and may forget to input decisions or key information into the system. This is why human observers are still necessary, but if the system can track and assess even half of the exercise play, it lightens the observer's burden considerably and allows him to concentrate on the finer points of the evaluation.

Assessment Criteria Authoring System

The assessment criteria authoring tool allows a domain subject matter expert (SME) to turn doctrine into assessment criteria. As with all instructional design, the effectiveness of the end product will depend largely on the author, so our authoring tool will be easy to use and give a clear view of how tasks, skills, and assessment criteria are related. It will allow the user to easily define when skills must be applied based on events in the simulation, how trainees can accomplish tasks and fulfill the skill assessment with actions in the simulation, and specify constraints, consequences, and standards. The tool will be usable by SMEs directly, so that assessment criteria is not filtered through data input personnel, reducing the opportunity for inaccuracies. The authoring system will allow for quick and easy doctrine updates.

A prototype of this authoring tool for AEAS-style skills training is described in Pigora, et al., (2002). Our objective system would extend this for question-based evaluation and the team metrics discussed in the previous section.

Presentation Methodology

As discussed previously, our presentation formats will include question-based tests for knowledge acquisition, simulation for an individual with simulated teammates for skill training, and simulation for several trainees working as a team for gaining experience to develop abilities. Each of these presentation methodologies will have variations that account for different learning styles and competency levels. For example, material presentation used with a question-based test could be mostly textual, provide diagrams for more visual learners, or small interactive modules for experimental learners. If a trainee is just beginning to learn the skills associated with his role in an individual simulation it may be beneficial for him to play with more proficient simulated teammates. Novice simulated teammates who make many mistakes could cause the trainee to take away incorrect information from a training experience. Multiple presentation methodologies will be used to ensure the training experience is tailored to both the information being taught and the learner.

Learning Management System

The purpose of our learning management system is to ensure a user's learning environment is suited not only to his proficiency level, but also to his learning style. According to Kolb (1984), a trainee is likely to oppose a learning environment that is not well suited to his personal learning style. In addition, a novice trainee will gain little from the experience of a simulated exercise because he doesn't have the basic knowledge to understand the experience. Information about each trainee is gathered and stored within the learning management system, and used to present the trainee with a training experience that will provide maximum benefit. This is accomplished, in part, through analysis of a trainee's performance with different presentation styles, including assessments and training frequency and duration.

A comprehensive AAR capability is incorporated into the learning management system. The AAR will display the automated assessment results, any notes or adjustments from human observers, and include a step by step walk through of the training exercise so all responses and actions can be reviewed. Reports can be generated and taken away as a soft- or hard-copy. All learning presentation methodologies use the same format AAR. Since all training experiences draw from the same assessment criteria, exercises can be compared side-by-side and trainees will clearly see their progression from learning the basics to gaining expertise. The AAR will also map negative assessment results to additional or remedial training material that will be available for the trainee to use immediately or in the future.

The learning management system will be entirely web-based. This will lessen a user's computer system requirements, provide more "on-demand" training opportunities, and allow geographically dispersed personnel to train together.

CONCLUSION

E-learning has been used effectively for training domain knowledge and individual skills. We have proposed a unified system that also trains abilities by allowing the trainee to practice applying skills in a team naturalistic decision making venue. Our objective system uses a single knowledge base to train and assess novices and experts through the three levels, and provides feedback in the team training environment by tracking aspects of team cognition and sources of team errors. In accordance with the Recognition Primed Decision model, the skills and team training exercises allow the trainee to play out their decisions and see the implications, building a pattern base via surrogate experience that enables them to better make decisions when real world challenges arise.

The success of e-learning in engaging the trainee will always depend in part on the content author. However, the interaction generated by training a team on a complex, relevant problem makes the exercise more engaging. By providing free-play environments to try out the answers and get feedback, we can replace expectations of boredom with interest in e-learning experiences.

ACKNOWLEDGEMENTS

The authors would like to thank the National Guard Bureau for supporting AEAS and the continuing work in command level emergency management team training.

REFERENCES

Bell, B., Johnston, J., Freeman, J., Rody, F. (2004) STRATA: DARWARS for Deployable, On-Demand Aircrew Training. *Proceedings of I/ITSEC 2004.*

Bloom, B. C. (Ed.). (1956). *Taxonomy of educational objectives: Handbook I, cognitive domain.* New York; Toronto: Longmans, Green: David McKay Company, Inc.

Elliott, Taryn (2005). Expert Decision-Making in Naturalistic Environments: A Summary of Research. Edinburgh South Australia, Australia: DSTO Systems Sciences Laboratory. Retrieved June 2, 2006, from http://www.dsto.defence.gov.au/publications/3812/DSTO-GD-0429.pdf

Greenagel, F.L. (2002). The Illusion of E-Learning: why we are missing out on the promise of technology. Phoenix. AZ: League for Innovation in the Community College, White papers. http://www.league.org/publication/whitepapers/0802.html

Klein, Gary (1998). *Sources of Power*, Cambridge, Massachusetts: MIT Press.

Klein, G. A., Orasanu, J., Calderwood, R., & Zsambok, C. E. (Eds.). (1993). *Decision making in action: Models and methods*. Norwood, NJ: Ablex.

Klein, Gary and Thordsen, Marvin (1990). A Cognative Model of Team Decision Making. *Proceedings of the 1990 Symposium on Command and Control Research*, 270-274.

Kolb, D. A. (1984). Ex*periential Learning*, Prentice-Hall.

Magerko, B., Wray, B., Holt, L., and Stensrud, B. (2005). Improving Interactive Training Through Individualized Content and Increased Engagement. *Proceedings of I/ITSEC 2005.*

Miller, T. E., Zsambok, C. E., & Klein, G. (1997). *Commander's cognitive demands in OOTW* (Final Report for National Defense University, Washington, D.C. prepared under Contract DADW49-98-P-0009). Fairborn, OH: Klein Associates Inc.

Pigora, M. A, Barshatzky, K., Kerrigan, M. and Murphy, H. (2002). A Task-Based Approach to Training, Evaluating, and Simulating WMD Response Behaviors. *Proceedings of I/ITSEC 2002.*

Salas, E., Sims, D.E., & Burke, C.S. (2004). Towards a taxonomy of teamwork: The core elements: Manuscript in preparation

Shanteau, J. (1992). Competence in experts: The role of task characteristics. *Organizational Behavior and Human Decision Processes* 53:252-266.

U.S. Department of Homeland Security (DHS) Office for Domestic Preparedness (2004). *Homeland Security Exercise and Evaluation Program Volume I: Overview and Doctrine*. Retrieved June 2, 2006 from http://www.ojp.usdoj.gov/odp/docs/HSEEPv1.pdf

Valenti, S., Cucchiarelli, A., & Panti, M. (2002). Computer Based Assessment Systems Evaluation via the ISO9126 Quality Model. *Journal of Information Technology Education, Volume 1 No. 3*.

Weil, S., Hussain, T., Diedrich, F., Ferguson, W., and MacMillan, J. (2004). Assessing Distributed Team Performance in DARWARS Training: Challenges and Methods. *Proceedings of I/ITSEC 2004.*

Welford, A. T. (1968). *Fundamentals of skill*. London: Methuen.