Title: **Benchmarked Experiential System for Training (BEST)**

Suggested Tracks:
Track 4: Cognitive and Social Issues
Track 8: C2 Modeling and Systems
Track 7: Network-Centric Experimentation and Applications

Authors:

Georgiy M. Levchuk, Ph.D.
Kevin Gildea, Ph.D.
Aptima Inc.,
12 Gill Street, Suite 1400
Woburn, MA 01801
Phone: 781-935-3966x267, x200
Fax: 781-935-4385
e-mail: georgiy@aptima.com

Jared Freeman, Ph.D.
Aptima Inc.,
1726 M Street, N.W., Suite 900
Washington, DC 20036
Phone: (202) 842-1548x316
Fax: (202) 842-2630
e-mail: freeman@aptima.com

Wayne Shebilske, Ph.D.
Ganesh Alakke
Shruti Narakesari
Wright State University
3640 Colonel Glenn Highway
Dayton, OH 45435

Correspondence:

Kevin Gildea, Ph.D.
Aptima Inc.
12 Gill Street, Suite 1400
Woburn, MA 01801
Phone: 781-935-3966x200
Fax: 781-935-4385
e-mail: kgildea@aptima.com

Abstract

We address the problem of optimizing instructional strategy for team command and control training exercises in simulator environments.

In the first phase of this work, we developed model-generated, near-optimal solutions to complex C2 scenarios, as well as animations and presentation techniques that supported their use as feedback. Experimental results demonstrated a reliable advantage for the group receiving this treatment, an advantage that could theoretically halve training time.

In the second phase of this work, we are combining three computational models to optimize the order of presentation of DDD C2 practice scenarios. The team developed (1) a communications model used to assess communications content; (2) an optimization agent that generates near-optimal solutions to the C2 scenarios and whose output is used as the standard against which measured human performance on the DDD is assessed; and (3) a POMDP model that recommends the next practice scenario (among many available) to accelerate team performance towards mastery of three competencies. Experimental validation is underway to validate this multi-model approach to optimizing team learning.

This work advances the science of training by developing models that explain and support team learning. In addition, this work is producing training content for air command and control teams, specifically those in AWACS and the Air Operations Center (AOC) Dynamic Targeting Cell (DTC).

Introduction

Expertise is a function of the amount of deliberate practice plus feedback. This simple formula derives from research concerning the genesis of expertise in exotic and everyday domains – from chess, music, and medicine to typing (Ericsson et al. 1993, Ericsson 2002, 2004), and it guides the design of training systems in three ways. First, training systems should be highly accessible, to promote frequent practice. Thus, training simulations should be delivered on the most portable, commonly available platform on which the targeted knowledge and skills can be trained. Second, practice scenarios in simulations should be systematically ordered and structured to address training objectives efficiently, that is, to promote deliberate practice. Third, performance assessments should be relevant to training objectives, and systematically delivered to train and maintain competencies through feedback.

In this paper, we address the problem of developing and delivering feedback concerning teamwork in ill-defined domains, and the challenge of structuring (ordering) scenarios so that they most rapidly advance teams towards expertise. We describe the Benchmarked Experiential System for Training (BEST) that implements solutions to these challenges, and experimental results concerning these solution strategies.

We tested the impact of the BEST solutions using a task that was exemplary of command and control in actual military settings. This degree of ecological validity improved the odds that the findings would generalize to the operational environment. The C2 'node' emulated was an Airborne Warning and Control System (AWACS) team. Cognitive task analyses of AWACS in operational settings (Fahey et al. 2000) guided the simulation's design.The AWACS team operates on board an E-3 Sentry aircraft (Elliott et al. 1999) equipped with a radar system that is capable of detecting airborne targets in excess of 200 miles. The crew usually has at least a Mission Crew Commander (MCC), Senior Director (SD), and several Air Weapons Controllers (AWC) (Formerly known as Weapons Directors or WD; Fahey et al. 2000). The AWCs have the primary responsibility for directing friendly aircraft. Each AWC may have oversight of a particular location, or 'lane,' or they may be assigned specific functions such as controlling High Value Assets (HVAs) such as intelligence gathering aircraft. The MCC and SD leave decisions in the AWC's hands and intervene only if there appears to be an oversight. Working under high workload, these teams make decisions in current, military operations that take or spare the lives of enemies as well as preserve the lives of our own warfighters and non-combatants.

Feedback Solution

The simulation required participants to defend a no-fly zone from enemy intrusion by taking one of three actions against enemy aircraft to maximize team scores: (a) attacking the enemy when it entered the no-fly zone (which yielded the training team a reward of 50 points less penalties for time spent in the no-fly zones); (b) preemptively attacking the enemy before it entered the no-fly zone (which was penalized because it violated the Rules of Engagement); or (c) ignoring the enemy, so that the target completed its intended path (which accumulated penalty points for the entire time the target spent in the no-fly zone). We developed near-optimal solutions to this task assignment and scheduling problem to use as feedback to teams. The optimization problem was tractable because: (a) each asset could strike only one enemy target before returning to base for reload/refuel; (b) each target could be attacked by only one asset (i.e. coordinating multiple assets to simultaneously attack a target was outside the scope of the problem); (c) complete target path parameters were known a priori.

The optimization algorithm has several interdependent phases.

- In Phase I, we find the allocation of targets to assets. Initially, each target is assigned to the closest asset whose capabilities are adequate to prosecute the target.
- In Phase II, we obtain a target sequence – the order in which the targets will be attacked.
- In Phase III, we find a task schedule – the exact times when the tasks will be prosecuted by assets, and specific actions taken by assets against the tasks (detection, identification, attack, etc.). The algorithm to find the optimal task times and associated launch schedule of the asset is based on a dynamic programming problem and accounts for influences of each task allocation on the execution of consecutive tasks in the task sequence.

We use a feedback among Phases I through III to iteratively improve the task schedule until a near-optimal solution is reached. When the schedule is obtained in Phase III and rewards and penalties for the schedules are calculated, we utilize an annealing approach to modify the allocation of the tasks to assets (Phase I) and then repeat Phases II and III to reliably raise the score of the schedule.

Feedback Experiment

The use of optimized solutions as feedback was evaluated in an experiment involving 120 graduate students in 30 teams of four. The teams executed defended two no-fly zones in a command and control task implemented on the Dynamic Distributed Decision Making (DDD) simulation (see Figure 1). Each team member, or decision maker (DM), controlled a base centered within one of four quadrants and four other assets, an AWACS, jet, helicopter, and tank. The bases were inside the corners of a green no-fly zone and outside the corners of a red critical no-fly zone. Six boxes in a blue report area displayed offensive scores (top three boxes) and defensive scores (bottom three boxes). The left, middle, and right boxes displayed scores for individual; groups for the North or South regions; and teams, the total score of all four DMs, respectively. The defensive scores dropped at the rate of one point per second for each enemy target in the green no-fly zone and two points per second for each enemy in the red critical no-fly zone. The offensive scores increased for every successful attack on an enemy target in the no-fly zones (individual + 5, group + 10, team + 25). The offensive scores decrease by 25 points each when an enemy target was destroyed outside the no-fly zones or when a friendly asset was destroyed anywhere. All these scores were displayed, but only the Team Defensive score, which started at 50,000, was used in the present analysis of team mission performance, because this score relates to the most important part of the mission, defending the no-fly zones.

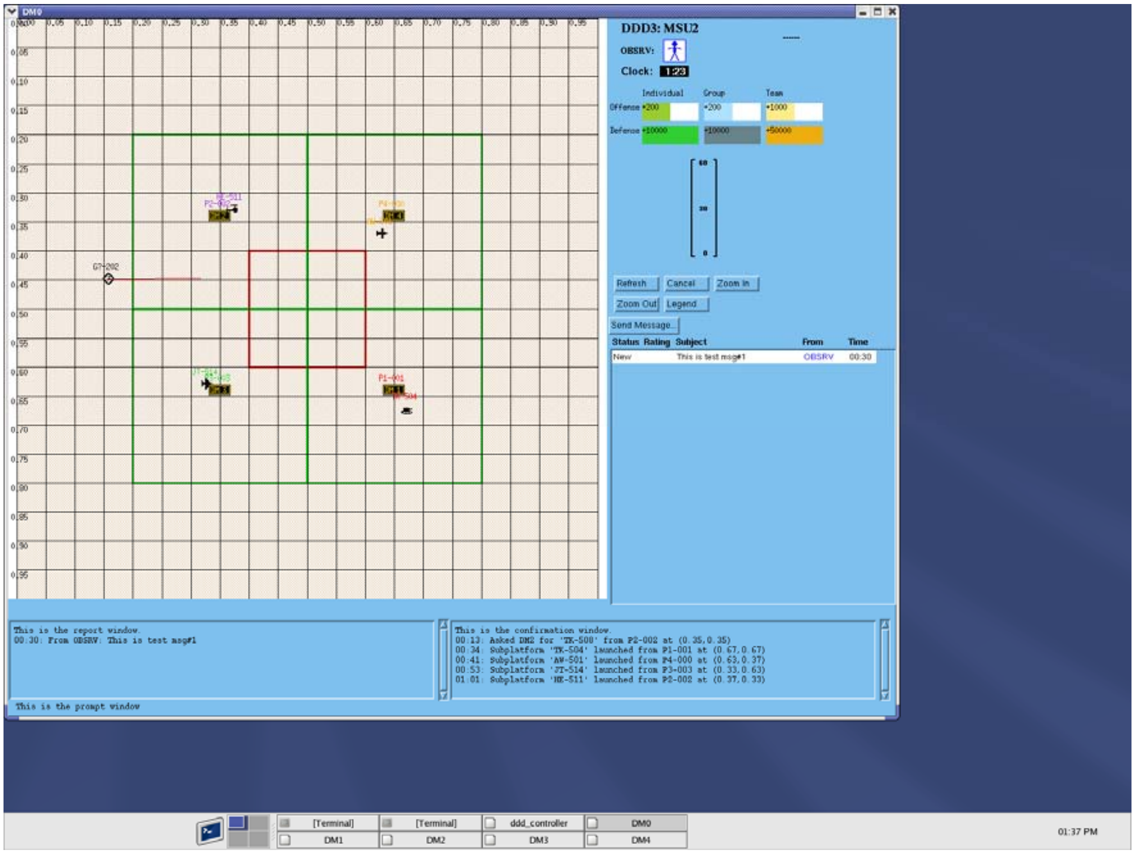

Figure 1: User interface to BEST scenarios on the DDD

The BEST experimental group received an animation of a near-optimal, BEST audio/video during debrief (described above). The solution presented was selected by a domain expert for its similarity to known expert solutions. The Control group received a list of general strategies. The two groups had equal time to review these debriefing materials and reflect on the previous mission during debriefing. The groups were treated the same in all other respects.

BEST feedback improved both the Team Offensive score and Team Defensive score. On the Team Offensive score, the BEST and Control groups were similar at baseline, but the BEST group performed significantly better throughout training ($F(1,28) = 6.47$, $p < .05$) with the final Offensive score being 1108 for the Control Group and 1177 for the BEST group.

A similar pattern is shown in Figure 2 for the Team defensive score on the mission assessments: 0 (Baseline) and 1, 2, and 3 for the BEST and Control groups. (We focus here on the scores for defending the no-fly zones, because it is the essential function of this type of mission.) The Team Defensive score results were analyzed with a split-plot ANOVA, with training protocol (BEST versus Control) as the between-participant variable and mission as the within-participant variable. There was a significant main effect of mission, $F(3, 84) = 389.75$, $p<.01$, and a significant interaction between protocol and mission, $F(3, 84) = 7.27$, $p<.05$. Planned comparisons showed that the BEST group performed about the same as the Control group on the baseline before training ($F(1, 29) = .18$, $p>.05$) and then performed consistently better on Assessment missions 1 ($F(1, 29) = 3.08$, $p<.05$), 2 ($F(1, 29) = 7.55$, $p<.01$), and 3 ($F(1, 29) = 4.21$, $p<.05$). The effect size in terms of the percentage of variance accounted for as measured by partial eta squared (η_p^2) for assessments 1, 2, and 3 were .10, .21, and .13. In practical terms, the best score an expert team with 100 hours of training has achieved is 48,000. If the trend in Figure 3 continued, this optimum would be reached at 7 trials for the BEST group (R2=.97) and at 14 trails for Control group (R2=.98).

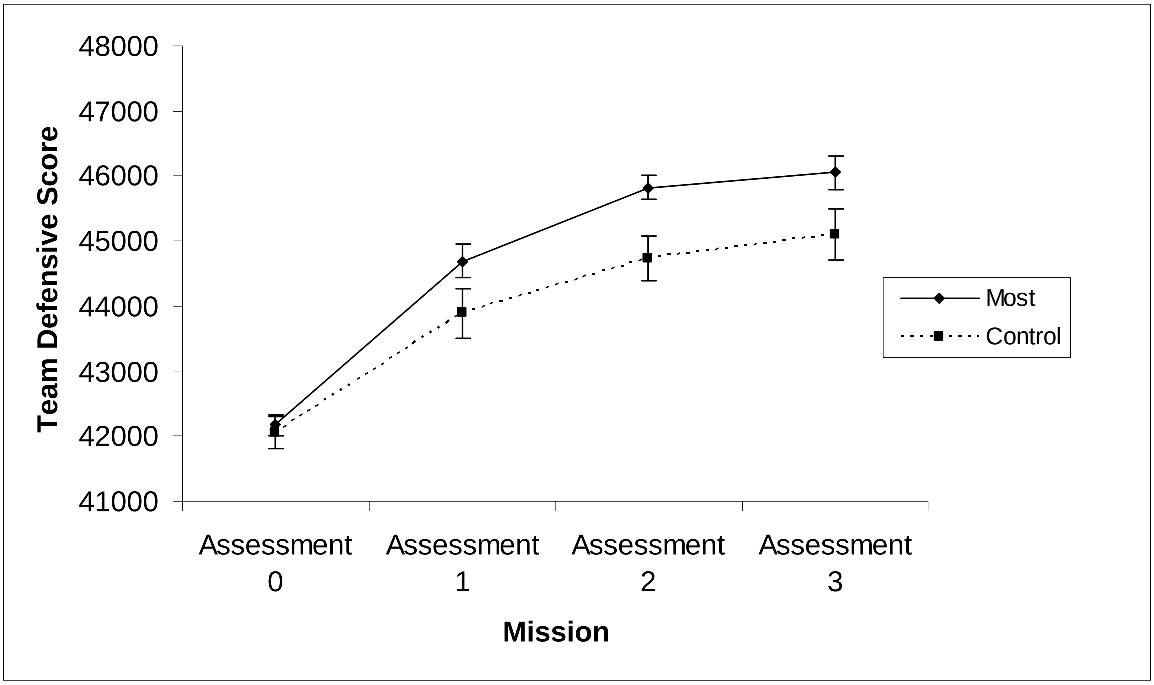

Figure 2: Performance effects of BEST group vs. control group

Scenario Ordering Solution

The effectiveness of training is a function, in part, of its structure. Here, we focus on the order of practice events, one aspect of structure. Three models have been developed to adapt the scenario order to the performance exhibited by a given team:

(1) A communications model is used to assess the content of written communications in the DDD environment;

(2) An optimization agent generates near-optimal solutions to the C2 scenarios. This model incorporates probabilistic resource to task allocation, probabilistic path planning, scenario based heuristics and logic, an objective function to maximize scores on a DTC scenario. Its output is used as the standard against which measured human performance on the DDD is assessed.

(3) A POMDP model recommends the next practice scenario (among many available) to accelerate team performance towards mastery of three competencies. Markov decision process models in general have proven to be effective in a variety of sequence and planning applications that involve elements of uncertainty in the process (Cassandra, 1998). The POMDP model takes input from the communications analysis model, other DDD measurement instrumentation, and the performance standards generated by the optimization agents. It applies an algorithm that represents the following to generate a table that enables the trainer to specify the next optimal scenario given its performance on the most recent scenario. Again, that model represents:

   a. a finite set of states (team expertise),

b. a finite set of actions (scenarios to train on),
c. a finite set of observations (performance/process measurements outcomes),
d. a state transition function (team expertise in relation to scenarios),
e. an observation function (how accurately measurement captures team expertise state), and
f. an immediate reward function (capturing scenario cost and expertise state score to be used for scenario selection)

These models function in a coordinated manner to enable trainers to capture performance measures in a largely automated manner and select the next scenario from a simple lookup table.

Scenario Ordering Experiment

Experimental validation is underway to validate this multi-model approach to optimizing team learning.

Conclusion

The present work is important in that it applies computational modeling techniques to the problem of improving instruction. Thus, it advances the science of training In addition, this work is producing training content for air command and control teams, specifically those in AWACS and the Air Operations Center (AOC) Dynamic Targeting Cell (DTC). This training content – 50 scenarios in all – is systematically scaled along several dimensions and implemented on a powerful and generally available simulator for team training and research.

References

CASSANDRA, A. (1998). A survey of POMDPS applications. American Association for Artificial Intelligence Symposium.

ELLIOTT, L.R., CARDENAS, R., and SCHIFLETT, S.G., 1999, Measurement of AWACS team performance in distributed mission scenarios. Available online at: HTTP://WWW.DODCCRP.ORG/1999CCRTS/PDF_FILES/TRACK_3/013ELLIO.PDF (accessed 28 August 2003).

ERICSSON, K.A., 2002, Attaining excellence through deliberate practice: Insights from the study of expert performance. In *The Pursuit of Excellence Through Education*, M. Ferrari (Ed.), (Mahwah, NJ: Lawrence Erlbaum Associates, 2002).

ERICSSON, K.A., 2004, Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine*, **79**, pp. S70-81.

ERICSSON, K.A., KRAMPE, R. TH., AND TESCH-ROMER, C., 1993, The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, **700**, p. 379 and p. 384.

FAHEY, R.P., ROWE, A.L., DUNLAP, K.L., AND DEBOOM, D.O., 2000, *Synthetic task design (1): Preliminary cognitive task analysis of AWACS weapons director teams.* Technical Report. Brooks AFB, TX: Armstrong Laboratory.