Designing a Distributed Team Training Experience to Balance Data Collection Needs With Natural Team Interaction

Randy Jensen [1], Grace Teo[2], Jim Lunsford [3], and Jeanine A. DeFalco[4]

[1]*Stottler Henke Associates, Inc., San Mateo, CA, USA*
[2]*Quantam Improvements Consulting, Orlando, FL, USA*
[3]*Decisive-Point, Kansas City, MO, USA*
[4]*US Army Combat Capabilities Development Command – Soldier Center, Orlando, FL, USA*

Abstract

This paper discusses tradeoffs in the design of a distributed trainer for U.S. Army war gaming exercises with automated team performance assessment. Automated assessment requires access to data reflecting team decisions and behavior. By the same token, the training experience must allow team interactions and decisions to flow naturally in the distributed setting, in order for participants to engage in meaningful teamwork. These factors can be at odds, when data collection methods make for less natural team interactions, or when the most natural experience presents obstacles to data collection. Tradeoffs evaluated for this application include questions involving the methods of communication during a war gaming exercise, the role of leaders as participants or facilitators, and the structured flow for team interactions and decisions.

Keywords

Team Tutoring, Teamwork, Team Performance Assessment, Distributed Training

1. Introduction

Instructional system design often involves tradeoffs between artificialities necessitated by practical elements of the learning environment, and the goals to make a realistic and compelling experience for learners. These tradeoffs play a significant role in the design of intelligent tutors, where automated performance assessment mechanisms seek to monitor decisions and behavior, in order to make inferences about mastery of different competencies. The challenge is to construct a training environment where learners perform as they do in the real-world, while the system can effectively trace decisions and behavior through instrumentation. Assessment requires mechanisms to capture sufficient data without imposing artificialities that skew performance or create a cumbersome training experience.

The tradeoffs are especially significant in team training applications, where naturalistic interactive processes such as communications are fundamental to teamwork. A number of these tradeoffs come into play with the design of a distributed team trainer for Army command-level war gaming. This is a team planning activity which is a structured component of the Military Decision Making Process (MDMP). War gaming is a process where multiple participants, representing a cross-section of warfighting functional areas, walk through proposed courses of action (COAs) to analyze them for suitability, feasibility, and acceptability. This requires significant teamwork, as the representatives of the different disciplines such as aviation, fire support, and intelligence contribute insights during the process of looking closer at the events in the COA, while also considering possible enemy reactions.

Proceedings of the Challenges and Advances in Team Tutoring Workshop during the Artificial Intelligence in Education Conference, June 15th, 2021, Virtual from Utrecht, The Netherlands
EMAIL: jensen@shai.com (A. 1); gteo@quantamimprovements.net (A. 2); jim@decisive-point.com (A. 3); jeanine.a.defalco.civ@army.mil (A. 4).

The training experience is designed to engage distributed teams in war gaming scenarios while their behaviors are assessed for several teamwork dimensions. The breakdown of dimensions applied to the domain of war gaming is adapted from common themes in the literature, with factors including leadership, supporting behaviors, information exchange, communication quality, team cognition, and team orientation [1-5]. Since participant behaviors are to be coded and scored on measures of these dimensions, it is critical that they are afforded opportunities to exhibit these indicative behaviors while war gaming in the training environment. Hence, the design must find a balance between the requirement to collect sufficient performance data to be used as indicators of teamwork in an exercise, and the requirement to provide an exercise experience where teams can effectively carry out war gaming in a way that is natural to the tasks.

2. Design Tradeoffs for Training Team War Gaming

Tradeoffs evaluated for this team training application include questions involving the means or channels of communication to be used during a war gaming exercise within the environment, the role of leaders as participants and/or facilitators, and the manner in which the flow of team interactions and decisions is structured during the exercises. These are each discussed in the following sections.

2.1. Tradeoffs in Team Communication Methods

In developing a team training experience, if data collection needs are given primary weight as a design factor, then in order to facilitate machine understanding of communication content, some constraints on communication methods may be utilized. This might mean relying on formatted or predefined messages, and limiting free-flowing team communications, which reduces the realism of the training experience. The opposite approach is to emphasize natural team interaction, and support or allow as much communication as possible, by any means available. There may inherently be hard constraints, as some forms of team communication simply cannot be easily simulated in a distributed virtual environment, such as the nonverbal cues used by a co-located team. Yet for any free-form communications such as voice or even chat, these can be readily supported in a virtual environment, but there are significant technical challenges in developing capabilities for automated understanding of the content. There have been numerous efforts to make communications assessment easier, but this is still an area of research and development. So the approach that emphasizes realistic team interaction has advantages for the training experience, but downsides due to the difficulty in either accessing or interpreting a significant data source for indicators of team performance to be used for assessment.

There are numerous possibilities for the design of communications methods in a distributed war gaming trainer, but for simplicity we mainly evaluated two competing approaches: one that requires all communications to be "in-game" using chat, and one that allows collaboration by other means even if those are not collected as data sources for assessment. These two approaches are described and contrasted in Table 1.

Table 1: Comparison of approaches for team communications in war gaming.

	"In-game" Mechanisms Only	Unrestricted Communication
Details	• Provide chat rooms • Require decisions to be expressed using overt input mechanisms • No other ways for participants to communicate	• Provide chat rooms • Require decisions to be expressed using overt input mechanisms • Allow other unmonitored communications (virtual, phone, same room)
Advantages	• Enforces collaboration using tools that can be fully monitored by the tutor	• Realistic, resembles real-world war gaming which is often conducted co-located • Conducive to team dynamics

		• Voice is not required, just allowed
Disadvantages	• Less realistic • Requires proficiency with user interface elements • Less engaging	• Voice and non-verbal communications are not monitored or assessed • Key process information for teamwork may be inaccessible

With both approaches above, a chat room is provided as one means of communication, and also specific decision inputs must be expressed overtly via user interface elements. For example, in the war gaming process, if the aviation lead chooses a route for a helicopter attack, and if the fire support lead agrees with the chosen route, both of these decisions must be expressed using direct inputs in the training environment. Then at a minimum the training system has information about the decisions, including key inputs across members within the team, which can be used to assess teamwork.

The two approaches differ in the amount of information available for assessment of the communications that relate to these decisions. Both collect chat data for assessment, but one allows additional unmonitored communications. In our example, did the fire support lead provide information about planned indirect fires, to influence the aviation lead's choice of routes? If such communications take place strictly over monitored channels such as chat, then this adds to the picture for assessment reasoning. If such communications may take place in completely unmonitored forms, then the assessment can neither analyze their content, nor even reason about the absence of communications, for example to draw conclusions about a lack of information exchange. However, the operational environment can impact behavior and team dynamics. If communications require extra effort with specific user interface elements, then team members may not volunteer information as readily as when they can simply verbalize it.

For the war gaming training application, subject matter expert discussions led to the conclusion that the design of the team communications experience should weigh realism more heavily than data collection. More specifically, artificial restraints on team communication methods would be a greater detriment to training objectives than a lack of data for assessment. So the design choice for this tradeoff is to plan for the approach where unrestricted communications are allowed, as long as key decision inputs are made directly in the training environment in a form that the training system can process.

One assumption built into this conclusion is that voice communications are difficult to assess. But although this remains an area of research, such capabilities may be feasible in the future. So the expectation that voice communications cannot be assessed is a near-term practical assumption, not permanent. Another assumption relates to the expected level of comfort for training participants using chat-based communication methods. On one hand, chat has become increasingly prevalent in the digital world, and in both the training and operations settings there are examples where chat has been the primary or only platform [6]. Yet, specifically for the command staff involved in war gaming, and the free-form, almost brainstorming nature of COA analysis interactions, chat can be sufficient but perhaps not ideal. The subject matter expert determination is that team dynamics for war gaming exercises would be most realistic if communication methods are not limited to chat. Thus despite the impact on data collection for assessment, the design emphasis is on the experiential benefits of conducting team war gaming exercises with communications as realistic as possible.

2.2. Tradeoffs in Designated Responsibilities for the Leader Role

When command staff go through the war gaming process in accord with MDMP practices, the leader customarily takes a directorial role, managing the walkthrough of events in the COA, soliciting inputs from team members, and ensuring that decisions are made. Often the commander is not present, so war gaming is led by an executive officer or chief of staff, who manages the conversation and the resulting decision products that will be reported back to the commander. In mapping this unique role to a distributed training setting, the designated leader is the natural person to inherit the analogous responsibility to facilitate the mechanics of the browser-based exercise with remote participants. For example, this involves making sure all parties are logged in and present, taking overt actions to step through COA events, and managing the level of detail for team analysis. In the real-world setting,

discussion topics are queued up conversationally, but in the distributed setting there are user interface actions required to manage the exercise for participants.

A standard design goal, especially for distributed training, is to try to minimize the degree to which participants need to be occupied with gaining familiarity with the training environment, instead of the training tasks themselves. Thus for the leader who will serve as exercise controller, it is important to provide as much support as possible to help manage the effective flow of exercises, ideally avoiding team confusion in working with the training environment. This is where the tradeoff arises, between supporting the leader's responsibility to control the mechanical elements of orchestrating the progression of a distributed exercise, and supporting the leader's function in directing COA analysis tasks and decisions in war gaming.

Table 2 below summarizes two different approaches for the nature of support to be provided to leaders in distributed war gaming exercises.

Table 2: Comparison of approaches for leader support in war gaming exercises.

	No Script or Cues for Leader	Script and Cues Provided to Leader
Details	• Leader minimum role is to manage progression through COA • Team members independently take tasks at each COA step, or may be assigned tasks if the leader chooses to • Leader is responsible for determining if a task needs to be assigned	• Leader receives cues throughout the scenario to trigger progression through the COA and assign tasks • Other team members receive tasking only from the leader • Leader is not required to recognize when tasks are needed, because cues are provided automatically
Advantages	• More opportunities for errors in leadership, information sharing, and other teamwork factors	• Avoid situations where the leader makes visible process errors • Exercise flow is regulated to proceed as intended
Disadvantages	• Unprepared leaders may have difficulty with the facilitator role without process cues • Exercise flow may go in unexpected directions or fail to proceed • Requires more custom user interface elements	• Less opportunities for teamwork errors due to the scripted flow, and thus less opportunities for assessment and feedback

In the first approach above where the leader is given no script or cues, the goal is to provide as much opportunity as possible for good or bad leader performance specifically with regard to the teamwork dimension relating to leadership. That is, at a minimum the leader has controls to synchronize the experience for all participants advancing through the COA, but all decisions about tasks are unprompted. In the analysis of a particular COA event such as a helicopter attack, if the aviation lead fails to take initiative and propose an attack route, then the leader is also not prompted by the system to assign that task. Thus, this is an opportunity for the leader to fail to demonstrate leadership, with the result being a less thorough war gaming process. However, this can also mean that intended parts of the exercise are not carried out.

In the second approach above, process cues are given to the leader in accord with a script for the war gaming exercise. The emphasis under this approach is on ensuring that the exercise unfolds correctly, and that the leader does not get distracted thinking about training environment processes when the focus should be on war gaming processes. Thus, a more rigid scripted approach is followed, where all tasking comes directly from the leader, using simple user interface controls to trigger the next decision point, the tasking to specific roles, and the changes to the shared user interfaces in the distributed setting so that specific roles can contribute their inputs.

To illustrate the contrast between the two approaches, we turn back to the example where the aviation lead needs to pick a route for a helicopter attack. In the first approach with no leader cues, the aviation lead needs to have a way to take the task of selecting a route based on initiative. So this requires a user interface action to start the task, which then opens the additional user interface elements for selecting the route –displaying several optional routes and an input tool. Or if the aviation lead fails to take this task, then the leader has a similar user interface control for assigning the task, again based on initiative. In the second approach, as soon as the exercise is advanced to the COA event involving the planned helicopter attack, the leader is given a one-step cue to trigger the tasking and display elements. The optional routes are displayed for all distributed participants, the aviation lead has toolbox controls for selecting a route, and a pre-scripted chat message is sent to all describing the tasking.

In this example, the unscripted approach provides more opportunities for teamwork errors. The aviation lead may fail to contribute a preferred route, the leader may fail to ensure that this input is coming, and other roles may similarly fail to contribute supporting information. However, the unscripted approach also creates more situations where participants may know what to do in terms of war gaming decisions, but not know which user interface controls to use for those inputs. It also places high visibility on the leader role as exercise controller. For situations where the leader is under-prepared for the distributed training event, or simply new to the environment, if the leader makes errors with the training environment processes then this can potentially degrade the training experience.

Based on subject matter expert input on this tradeoff, the use of scripting and cues for leaders is the preferred initial approach for several reasons. The scripted approach enables a more automated flow of events and tasking, with the goal of minimizing busy user interface elements and reducing confusion about how to express inputs in the exercise environment. The leader role is the one most significantly impacted by this approach, because the cues mean that leaders are not required to recognize when tasking has been overlooked, and this means there are fewer opportunities to assess the teamwork factor of leadership. However, for other roles, their tasking can appear to come from the leader under either approach, and they still have many opportunities for teamwork behaviors – sharing information, and supporting each other across functional areas. In addition, although a script and cues may constitute a degree of coaching or a training intervention for the leader and indirectly for the team by extension, the initial goal is not to evaluate training interventions in isolation for this research effort. Instead, the intent is to assess the teamwork competencies when a team trains in war gaming. If the leader is given a script and cues, this helps ensure that the data collected reflect behaviors associated with a level of war gaming training. Nevertheless, a modified approach to be explored in the future is to design a training sequence that transitions from a higher level of leader support as initial scaffolding, to a reduced level of support where cues are taken away as participants and especially the leader gain familiarity with the training environment.

2.3. Tradeoffs in Sequencing Team Member Inputs

This tradeoff is concerned with how the exercise flow is designed, and the collaborative process established for the command staff team performing war gaming. Especially in a distributed setting, collaboration can be hindered when there is uncertainty as to what other team members are doing. So this leads to a design question regarding the choice of a more parallel versus serial exercise workflow, a distinction that refers to how different team members make decisions and contribute inputs during the process. War gaming is conducted as a team, to refine and synchronize plans for each event in a COA. When the staff consider a COA event such as a helicopter attack, there may be several inputs or decisions related to that event, to be contributed by different roles. If participants work in parallel, they are less aware of what others are working on (if anything), but there is a time saving element in working independently. If participants work serially, the entire team focuses on a single particular topic or decision at the same time. All team members know the question at hand, and have a greater awareness of what others are thinking about, but also this means that more thinking takes place “on stage” with a certain amount of added pressure for deliberative processes. Table 3 below describes two contrasting approaches for structuring the collaborative process, using either a more parallel or serial workflow.

Table 3: Comparison of approaches for sequencing team inputs in war gaming exercises.

	Parallel and Independent	Serial and Fully Collaborative
Details	• Participants in different roles sign up for tasks so others know they will work on them • Different participants may be working on different inputs simultaneously • Inputs and decisions are worked out independently before submitting them	• Leader assigns tasks by role; all participants focus on the same task or topic simultaneously • Inputs are made "on stage" in a collaborative process with the entire team seeing decisions as they are made
Advantages	• More parallel independent work, especially when calculations are needed • More discrete decisions for assessment	• Resembles the collaborative process in real-world war gaming • All roles see decision-making process firsthand • Heightened collaboration on each task potentially leads to more opportunities for supportive behavior in information sharing
Disadvantages	• Less fluid exercise flow • Less engaging when distributed roles are working independently • Potentially more opaque decision-making	• Reduced situations where tasks need to be offloaded, which means less opportunities for backup behaviors • Potentially more time consuming

On one hand, independent work by individuals in parallel would seem to yield more structured opportunities to identify the individual versus team factors in decisions. If participants in different roles need to take overt actions to assume tasks, then there are also opportunities for backup behaviors when different roles or the leader see that there are needed tasks that have not been taken. This can happen regularly with the command staff engaged in war gaming, as officers serving in one role may have past experience in other warfighting functional areas, and therefore readily exhibit backup behaviors. In a training environment constructed to collect explicit data points for who takes what tasks and when, there are greater opportunities to use these data for teamwork assessment. However, parallel work makes for a more opaque decision-making process, especially in a distributed trainer where participants may be in remote locations.

According to subject matter expert input, real-world war gaming is conducted more as a group discussion. Although the process is managed, in the sense that the leader initiates the discussion for each topic, inputs are more free-flowing as individual roles take the floor and step through their decision-making with an open opportunity for inputs from other roles. This is likely to be more engaging even for roles not directly involved in a COA event or topic, as they can see the decisions unfold along with the underlying reasoning, which may be less apparent if the workflow involves more parallel independent decisions. For these reasons, the initial distributed trainer design is planned around a serial, fully collaborative workflow, with its resemblance to real-world war gaming. Although the data stream under this workflow likely has less clear boundaries between collaborative contributions from different roles during discussions, and thus there are less opportunities to assess teamwork factors such as supporting behavior, the actual teamwork under this exercise design is likely to be more effective. A related benefit to the fully collaborative team workflow in real-world war gaming is that it likely produces more shared awareness of the battlefield decision points and synchronized planning objectives. These effects can be measured during the exercise with probing questions about team cognition.

3. Summary and Conclusion

This paper discusses three different kinds of tradeoffs considered for the design of war gaming exercises to be conducted in a distributed team trainer with automated teamwork performance assessment. All the tradeoffs directly relate to the kind of experience participants have while working as a team, including how they communicate, what responsibilities the leader is given in controlling the exercise, and how tasks are performed individually or collectively during the exercise. A common theme in the different tradeoffs is the competition between the data collection needs for automated assessment to be effective, and the inherent need to have a realistic exercise experience where the team can perform naturally. Ultimately the greater emphasis is placed on realism, that is, creating or allowing natural team interactions with each other and with the system.

Judgment calls about the necessary level of realism are mainly based on subject matter experts, but still partly weighed against practical considerations for implementation. For example, regarding the level of system support provided to leaders for exercise control, the decision to provide cues is motivated more by practical concerns in the distributed training environment than by realism per se. The plan is to develop a fully operational training prototype that applies the initial design decisions. This will be used to test assumptions regarding the suitability and acceptability of the training experience, by collecting feedback from the end user community and related stakeholders.

4. Acknowledgments

The research reported in this document was performed in connection with contract number W912CG-20- C-0014 with the U.S. Army. The views presented in this paper are those of the authors and should not be interpreted as presenting official positions, either expressed or implied, of the U.S. Government.

5. References

[1] J.H. Johnston, J.A. Cannon-bowers, E. Salas, Tactical Decision Making Under Stress (TADMUS): Mapping a program of research to a real world incident—The USS Vincennes (1998).

[2] S.W. Kozlowski, J.A. Grand, S.K. Baard, M. Pearce, Teams, teamwork, and team effectiveness: Implications for human systems integration (2015).

[3] S.L. Marlow, C.N. Lacerenza, J. Paoletti, C.S. Burke, E. Salas, Does team communication represent a one-size-fits-all approach?: A meta-analysis of team communication and performance. In Organizational Behavior and Human Decision Processes, 144 (2018):145–170. https://doi.org/10.1016/j.obhdp.2017.08.001

[4] E. Salas, D.E. Sims, C.S. Burke, Is there a “big five” in teamwork? In: Small Group Research, 36(5) (2005): 555–599.

[5] R.A. Sottilare, C.S. Burke, E. Salas, A.M. Sinatra, J.H. Johnston, S.B. Gilbert, Designing Adaptive Instruction for Teams: A Meta-Analysis. In International Journal of Artificial Intelligence in Education, 28(2) (2018): 225–264. https://doi.org/10.1007/s40593-017-0146-z

[6] S. Ramachandran, R. Jensen, O. Bascara, T. Carpenter, T. Denning, S. Sucillon, After Action Review Tools For Team Training with Chat Communications. In Proceedings of the Industry/Interservice, Training, Simulation & Education Conference (I/ITSEC 2009) (2009).