Team Training for Collaborative Cross-Functional Problem-Solving in Wargaming Exercises

Randy Jensen
Stottler Henke Associates, Inc.
San Mateo, CA
jensen@shai.com

Grace Teo
Quantum Improvements Consulting
Orlando, FL
gteo@quantumimprovements.net

Lisa Townsend
U.S. Army CCDC-SC
Orlando, FL
lisa.n.townsend2.civ@army.mil

ABSTRACT

Course of action (COA) development and wargaming are critical stages in the U.S. Army's military decision-making process (MDMP), where command staff at battalion and higher echelons analyze COAs in detail, considering actions, enemy reactions, and counteractions. Teamwork behaviors involving coordination across warfighting functional areas are essential to wargaming effectiveness and therefore central learning objectives in training. Full-scale wargaming exercises involve teams-of-teams, as different functional leads and their subordinate staff make decisions and projections relevant to major COA events, to produce a detailed, synchronized operational matrix. In order for a training audience to work effectively in full-scale wargaming, there is a need for preparatory collaborative exercises where they can gain greater awareness as to their specific roles and contributions, and those of other team members. This paper describes a distributed training tool developed to prepare staffs for full-scale wargaming, with a browser-based synchronous team dimensional familiarization exercise guided by an instructor. The exercise is presented as an integrated wargaming vignette, where command staff participants are prompted to review COA events and consider the role and impact of different warfighting functional areas. Since the exercise focuses on teamwork, much of the emphasis is on participants identifying relevant questions to be considered in the decision process, without the additional overhead of determining scenario-specific answers. Team performance assessment is captured via teamwork markers associated with an assessment model organized around team dimensions tailored for wargaming. Markers are created both by instructors and automated system rules during the exercise, for later incorporation into a debriefing for team self-correction. Markers are tagged with salient team constructs, qualitative assessments, and annotations about the flow of knowledge across warfighting functions. This paper presents initial training tool feedback from experienced instructors, as well as further discussion of the tradeoffs in team training for wargaming and other collaborative problem-solving domains.

ABOUT THE AUTHORS

Randy Jensen has been a project manager for over 20 years at Stottler Henke Associates, Inc., in San Mateo, California, where he has developed intelligent tutoring systems (ITSs) for domains involving complex decision-making. Examples include trainers for air attack planning, unmanned vehicle command and control, small unit tactics, combined arms team exercises, information systems troubleshooting, satellite scheduling, and a current effort for course of action development and wargaming at battalion and higher echelons. Randy has a B.S. with honors from Stanford University.

Grace Teo, Ph.D. is a Senior Research Psychologist at Quantum Improvements Consulting. Much of Grace's research involves understanding and improving human performance in various contexts and developing assessments to understand how we work with different technologies, and in teams. Her work involves research into the use of adaptive systems and emerging technologies in training, human-agent interaction, and human-centered applications of artificial intelligence (AI). Other research interests include automation, decision making processes and measures, vigilance performance, and individual differences.

Ms. Lisa Townsend is a Psychologist who has been working at STTC (SED/CCDC SC) in Orlando, FL since August 2022. Prior to STTC, she spent 27 years as a Research Psychologist at the Naval Air Warfare Center Training Systems Division (NAWCTSD) in Orlando, FL. She has a Master of Science in Industrial/Organizational Psychology and a Bachelor of Arts in Psychology, both from the University of Central Florida (UCF). She has worked on many diverse teams including those within Research and Development, Technology Transfer, Instructional Systems Design, and

Human Systems Integration. Ms. Townsend's areas of expertise involve team training related research, Front End Analyses (FEAs), Training Systems Analyses (TSAs), Instructional Systems Design (ISD), Training Effectiveness Evaluations (TEEs), and the development of training and organization related metrics. Her efforts in these areas have spanned across Services and platforms.

Team Training for Collaborative Cross-Functional Problem-Solving in Wargaming Exercises

Randy Jensen
Stottler Henke Associates, Inc.
San Mateo, CA
jensen@shai.com

Grace Teo
Quantum Improvements Consulting
Orlando, FL
gteo@quantumimprovements.net

Lisa Townsend
U.S. Army CCDC-SC
Orlando, FL
lisa.n.townsend2.civ@army.mil

INTRODUCTION

Course of action (COA) development and wargaming are critical stages in the U.S. Army's military decision-making process (MDMP), where command staff at battalion and higher echelons analyze COAs in detail, considering actions, enemy reactions, and counteractions. Teamwork behaviors involving coordination across warfighting functional areas are essential to wargaming effectiveness and therefore central learning objectives in training. Full-scale wargaming exercises involve teams-of-teams, as different functional leads and their subordinate staff make decisions and projections relevant to major COA events, to produce a detailed, synchronized operational matrix. In order for a training audience to work effectively in full-scale wargaming, there is a need for preparatory collaborative exercises where they can gain greater awareness as to their specific roles and contributions, and those of other team members. This paper describes a distributed training tool developed to prepare staffs for full-scale wargaming, with a browser-based synchronous team dimensional familiarization exercise guided by an instructor.

The exercise is presented to the training audience as an integrated wargaming vignette, where command staff participants are prompted to review COA events and consider the role and impact of different warfighting functional areas. There is a distinction between taskwork, or the task-specific behaviors related to performing the task at hand, and teamwork, or the set of behaviors that facilitate the coordinated functioning of the team itself (Cannon-Bowers et al., 1995). Since this exercise focuses on teamwork, much of the emphasis is on team members identifying relevant questions to be considered in the decision process, without the additional overhead of determining scenario-specific answers. For example, in full-scale wargaming, an assessment of maneuver routes may require the command staff leads to work with their own subordinate staff to produce estimates for considerations of time, distance, terrain, logistics, and a host of other factors that combine their own operational expertise with scenario-specific information. In the more limited scope of a preparatory exercise, the focus is on identifying these factors (the questions to ask), but not on producing the estimates. This allows a focused group of staff leads to collaborate in the exercise and gain practice with cross-functional coordination, in a shorter timeframe than full-scale wargaming.

Although it is an objective for the exercise and training tool to make use of automated adaptive instructional methods, the training use case still fundamentally relies on the role of the human instructor. So automated methods are designed to help the instructor, and the tool uses a hybrid approach to assessment. The instructor plays a significant role in identifying noteworthy examples of teamwork, but automated rules help in certain conditions. Team assessments are captured via teamwork markers, which are created both by instructors and automated system rules during the exercise, for later incorporation into a debriefing for team self-correction. The wargaming preparation exercise emphasizes assessment of team processes. Where team performance relates to team task or status outcomes, team processes relate to interactions within functions, to others, or from other team members. In a wargaming context, team performance examples include planning decisions for events in the COA (e.g., a choice of a route or position, unit tasking, synchronization, methods of attack, etc.), and team processes can include any of the dynamics that lead to these collective decision outcomes (Jensen et al., 2021). For example, the collaboration that takes place while deciding upon a route is more of an emphasis for assessment than a measure of the tactical merits of the route chosen. Markers are tagged with salient team constructs, such as supporting behavior, team cognition, or information exchange, referencing a model tailored for wargaming (Teo et al, 2021). Supporting information associated with markers also includes qualitative assessments, and annotations about the flow of knowledge across warfighting functions.

This paper discusses background considerations associated with teamwork in wargaming, as a preface to an overview of the design of the wargaming preparation exercise and tool. This is followed by a review of initial training tool feedback from experienced instructors, and takeaways for this and other collaborative problem-solving domains.

BACKGROUND: WARGAMING AS TEAM COLLABORATIVE PROBLEM-SOLVING

Task Structure

In characterizing team tasks, researchers distinguish between task and outcome interdependence, pertaining to the arrangements or structure of the team task. Task interdependence is high when team members have different access to critical resources and information, requiring them to rely on each other and coordinate their workflows to accomplish task goals. Individual tasks that contribute to overarching team-based tasks require interdependence coordination (i.e., the extent to which successful performance and meeting objectives relies on team members working together). Outcome interdependence is high when task goals are framed at the group level and when feedback or rewards are allocated at the team level rather than the individual level (Courtright et al., 2015). As a team task, wargaming is structured to have high task and outcome interdependence. The command staff team, comprising staff members who possess expertise in their respective warfighting function, must collaborate and coordinate their efforts toward the common goal of formulating a viable COA that is feasible, acceptable, and suitable. Effective collaboration and coordination are the hallmark of staff teams that produce efficient and effective decisions, though this can be especially challenging for novice command staff. As characterized by instructors, many in the incoming training audience have experience in their warfighting functional area but not as members of a command staff at higher echelons. Unlike their regular day-to-day responsibilities within their functional elements, collective planning activities like wargaming require these staff leads to serve as liaisons to the headquarters, a role in which novice command staff members have limited experience. In such a role, the command staff member not only needs to know the general purview of each warfighting function, but they must also comprehend and exploit the synergies among and across functions at the higher echelons.

Cannon-Bowers and Bowers (2011) proposed four categories of team tasks based on interdependence:

- Pooled interdependence (group output is the sum of individual output; e.g., sales teams)
- Sequential interdependence (group output is a sequence of individual output; e.g., assembly lines)
- Reciprocal interdependence (group output is an interaction between two members; e.g., command-and-control teams)
- Team interdependence (group output is an interaction among all members; e.g., self-managed work teams)

Although task structure "sets the stage" for teamwork behaviors to emerge, and specific configurations can encourage teamwork, they do not guarantee it. Task structure alone is insufficient in facilitating the teamwork essential in wargaming. For instance, even though rewards are associated with the team's output more than a simple sum of individual efforts, it still may be the case that the team's output is produced with a disproportionate contribution from a single member or subset (Courtright et al., 2015). In a command staff team, the team outcome may be more reflective of the inputs of a dominant member (individual sequential interdependence) instead of the overarching goal of being greater than the sum of individual members' inputs (team interdependence).

Nature of the Task

Many team tasks high on task and outcome interdependence, such as tasks undertaken by surgical or production teams, involve members collaborating to produce a tangible outcome. In these tasks, members contribute their specialized skills and resources, and work together towards the common goal. The work at both the individual and team levels is typically well-defined by members' areas of expertise or by team hierarchy. There is usually some order of task execution that helps outline expectations for how the work would proceed. This is not the case for wargaming. Wargaming is a team problem-solving task where the command staff team collaborates to "solve the problem" of uncertainty in battle by providing clarity to the Commander through their analyses. Clarity is obtained by addressing questions, such as: What is the mission? What is the constitution of friendly forces? What is the nature of the threat? What is the composition and disposition of the friendly forces relative to that of the threat? What are the strengths and weaknesses of the friendly forces relative to that of the threat? What is the terrain/weather/season? How do friendly forces synchronize their actions to achieve the COA objectives? The team's work is mainly cognitive and places a high demand on critical thinking at the individual and team levels. This critical thinking includes identifying the gist amidst volumes of material, synthesizing disparate information, making appropriate generalizations, and deciding when to seek information based on its value and costs (Fischer et al., 2008). Critical thinking also involves metacognition skills such as being able to identify assumptions, recognize cognitive bias in one's thinking and that of others, and distinguishing what is given in a particular situation (i.e., the grounds) from what is inferred or decided

upon in that situation (i.e., the conclusions) (Cianciolo & Sanders, 2006; Cohen & Thompson, 2001). Teams today are responsible for accomplishing cognitively complex tasks, which require them to plan, decide, remember, make decisions, solve problems, and generally think as an integrated unit (Cooke et al., 2008).

In their discussions, the staff team seldom knows how the analyses would unfold or when different staff members would be required to contribute since any one idea can lead to multiple other ideas given the unique perspectives of staff members, and the direction of the analysis can change with the type of questions asked or how questions are framed and addressed. That the team's conceptual analyses can be derailed or brought back on track with a single idea or question underscores the value of asking the right questions and the significance of the leader's or facilitator's role. To ask the right questions and address questions posed, each staff team member must draw on (i) what they understand of the Commander's intent and mission, (ii) what they know of their own warfighting function and others, (iii) what they need from others to be more effective, (iv) and what they must provide to others for them to be effective. In addition to exchanging their function-specific information and skills, they also contribute by feeding off others' questions and responses and exhibiting supporting and team orientation behaviors.

Becoming an Effective Team Quickly

Apart from being composed of staff members with various experiences, backgrounds, and training, command staff teams rarely remain intact for long, as members frequently rotate in and out of the team. Since shared understanding from common experiences and familiarity among team members are common qualities in effective teams (Gibson & Cohen, 2003), this arrangement can tend to produce command staff teams that may not be effective at the outset. Given the high stakes associated with warfighting preparations, however, these teams must be effective from the outset. To facilitate team development, staff members should be provided with multiple opportunities to work together with others on a variety of scenarios that are accompanied by guided learning experiences and developmental feedback (Mathieu et al., 2000). Such learning experiences should involve preparatory training that emphasizes understanding the interrelatedness of the warfighting functions, especially at high echelons, which impact the quality of the questions and depth of analysis in the subsequent wargaming. Providing a diverse, often ad hoc, command staff team with guidance and tools to inform the after action review (AAR) discussion is imperative in becoming an effective team more quickly. AAR interventions encourage reflection and self-discovery, target opportunities for improvement, and as a result, improve the quality of experiential learning, which thus improves team inputs, processes, and outcomes (Tannenbaum & Cerasoli, 2013). Research on and application of AARs have increased in recent decades. A meta-analysis conducted by Tannenbaum and Cerasoli (2013) found that AARs resulted in an average 25% improvement compared with control conditions (d=0.66). AARs can be used to address and impact team current and future outputs (Shuffler et al., 2018), which can be vastly useful and mission critical for training command staff teams.

EXERCISE DESIGN

Exercise Structure

The goals of preparatory training and team familiarization are central to shaping the design and scope of the exercise. In contrast to full-day or multi-day wargaming, this exercise is designed for a running time of approximately two hours, including the time for feedback and team reflection in AAR. This compact length is intended to optimize staff participants' time in gaining exposure to multiple vignettes and scenarios for effective team practice. For similar reasons, a design choice was made to implement the exercise platform as a distributed, browser-based synchronous team trainer. The purpose is to allow participants to be either remote or co-located, affording flexibility so that exercises *can* be conducted as homework away from a schoolhouse, or in conditions where not all participants are available in person. This flexibility is partly motivated by experiences during the pandemic, when it became much more difficult to conduct team training events in-person or with existing videoconference tools alone. On the other hand, remote exercises also add artificial constraints on the free flow of teamwork behaviors, so this design choice relates to one of the tradeoffs explored when initial feedback on the exercise was gathered.

Participants

The wargaming preparation exercise is designed to be led by an instructor, who directs and manages the flow of events, and also assumes the role of the Lead planner on the command staff. The role of Lead planner has an analogous

function in real-world wargaming, managing the walkthrough of critical events to be discussed. Exercise participants are each assigned to one of six command staff roles representing the leads from different key warfighting functions:

- Intel
- Movement & Maneuver (MM)
- Indirect Fire (Fires)
- Aviation (AV)
- Protection (PRO)
- Sustainment (SUS)

Exercise Activities and Flow

An exercise begins with the Lead / Instructor directing participants to review pre-briefing materials about the operational scenario, such as a tactical map, mission statement, and enemy order of battle. These are static materials available in the main reference panel throughout the exercise; although some scenarios may use several static map images, there is no underlying simulation. The exercise environment includes a status panel showing the roles that are logged in, and a shared message panel which is where most collaboration and teamwork behaviors take place. Participants using the message panel may either reply inline on existing threaded discussions, or add new messages which are treated as new threads. Participants may communicate with each other by other external means (for example, if training is conducted with an accompanying video conference or simply in a co-located setting), but the training application has no record of those interactions. In cases where offline discussions occur, a standard practice would be to encourage participants to use the message panel to post the results of any significant external discussions, for the record. Figure 1 below shows the Lead / Instructor screen, which includes a panel unique to this role (lower left, with a list of Topics and a Prompt field). For all other participants this panel is not visible, but the other panels are the same.

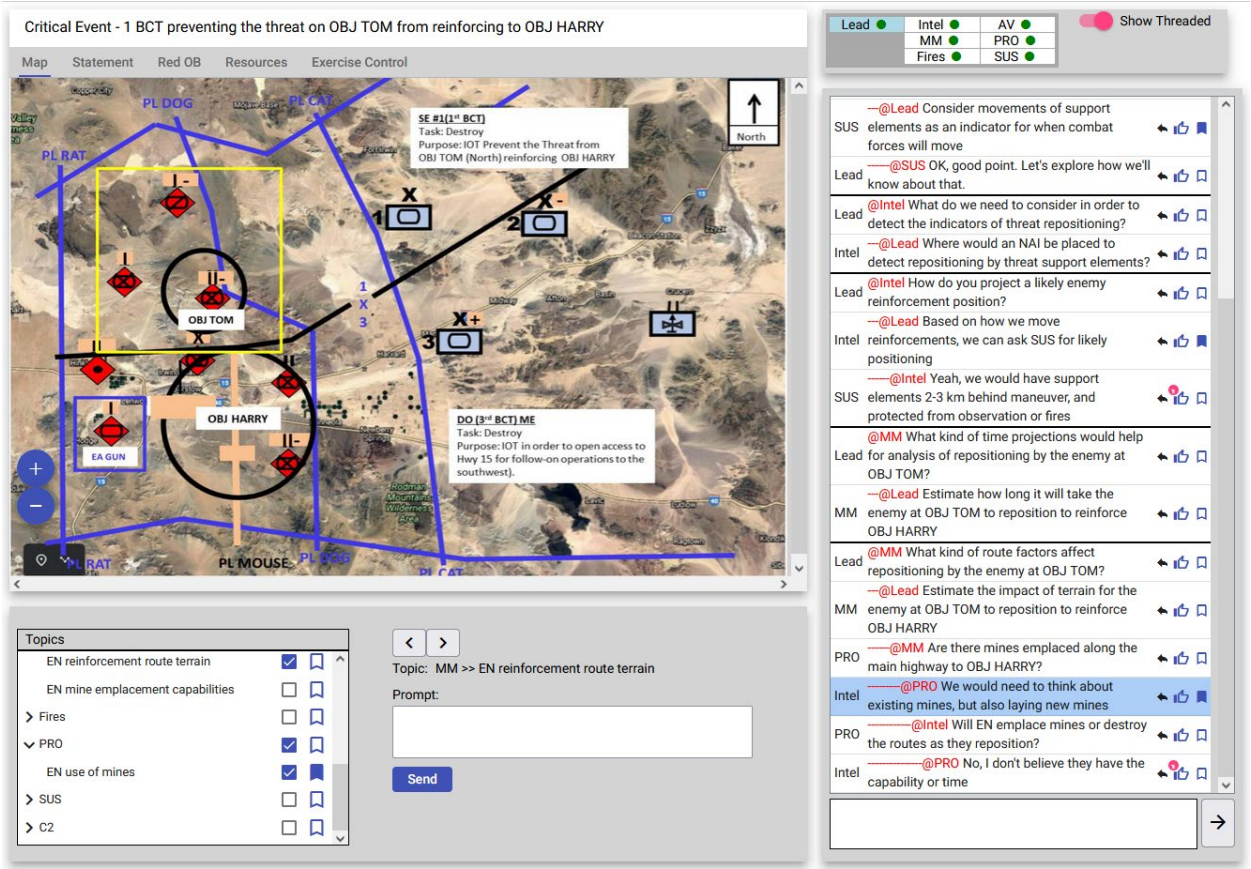

Figure 1. Wargaming Preparation Exercise Environment – Lead / Instructor Screen

The Lead / Instructor manages the exercise according to a predefined list of Topics. The Topics correspond to a mental model checklist that Instructors have used on paper in the past, to track progress and coverage of key concepts in wargaming-related exercises. Topics are organized by warfighting function, as the process of COA analysis naturally involves groupings of considerations related to each functional area. Each Topic has a predefined Prompt that the Lead / Instructor can directly post to the message panel to initiate staff discussion, or modify before posting. The Topics are only visible to the Lead / Instructor, and not to the training audience. During the exercise, noteworthy examples of teamwork (either positive or negative) are identified with Markers. Markers can be created either by the Lead / Instructor or by automated rules in the system. When created, Markers are associated with a specific message, Topic, or a more general observation, to capture significant examples of teamwork from the exercise for further discussion in AAR.

As participants type messages, a selection of auto-complete options is available as they type, and they can pick one of these options or just continue to type free text. Since the auto-completes are collected in a predefined library with metadata for their meaning and relevance to Topics, this is one way for the system to understand messages coming from participants. In such cases, automated system rules can assist the Lead / Instructor with notations about messages posted by staff members, and what they reflect in terms of Topics and / or teamwork behaviors.

In the exercise snapshot above, discussion threads are separated by horizontal dividers. The most recent discussion starts with a Prompt from the Lead / Instructor, addressed to Maneuver (MM), to identify considerations for enemy repositioning routes. Prompts are often written in an open-ended way ("What kind of route factors…") to avoid being overly suggestive of the Topic that participants are expected to consider. In this case, the Maneuver role responds using an auto-complete that expresses an appropriate consideration about terrain impacts along the routes from OBJ TOM to OBJ HARRY ("Estimate the impact of terrain…"). Since the system recognizes this auto-complete input as a suitable answer, it makes a notation that the corresponding Topic has been covered (in the Lead / Instructor's Topics list above, "EN reinforcement route terrain" is checked). The Lead / Instructor may also take similar actions with the Topic checkboxes, or undo system actions. In addition to the automated rules that note when Topics are covered, there are also rules that generate Markers when (auto-complete) inputs are contributed in ways that suggest good teamwork, such as supporting behavior from secondary roles.

Continuing with this example, the team discussion thread goes further, as the Protection (PRO) role brings up another key consideration concerning enemy mine emplacement along routes. In doing so, the ensuing comments between Protection and Intel cover a Topic ("EN use of mines") on the Lead / Instructor's list of planned Topics. This not only shows good cross-functional collaboration, but also good initiative, as the Lead / Instructor had not yet prompted the staff to consider concerns related to enemy mines. Since the staff anticipated this Topic on their own, the Lead / Instructor elects to create a Marker for this interaction, as an example of good teamwork behaviors. The Lead / Instructor selects a key message to associate the Marker with (Intel: "@PRO We would need to think about existing mines…"), and clicks on the bookmark / ribbon icon on that row to bring up a Marker dialog window as shown in Figure 2 below.

Since the Markers serve as an instructional record for reference in AAR, the Lead / Instructor uses the Marker dialog window to add several kinds of supplemental information. In this case, the Lead / Instructor makes a qualitative assessment that this example is Above expectation, and a High priority for AAR. The Knowledge of Functions checkboxes correspond to an existing team training AAR practice described by instructors, where each participant explicitly reflects on what they learned about different warfighting functions and roles within the team. In AAR, participants are encouraged to think about knowledge gained WITHIN their own function, knowledge they need to provide TO other functions, and knowledge they need FROM others. The checkboxes allow the Lead / Instructor to make a note of the kind of cross-functional interactions to highlight with a Marker. These are often reciprocal relationships (when one role needs to provide knowledge TO another, that other role needs it FROM the first), but even with reciprocal information sharing, often it may be one side of the communication exchange that has more of a new discovery about the interrelationship between roles. So the Lead / Instructor may check more than one box if desired, or whichever aspect of functional knowledge exchange is the intended emphasis or focus for the Marker. In the example below, the Lead / Instructor also adds a note about the discussion between Intel and Protection, and tags this Marker with the Team Cognition teamwork dimension relating to shared mental models and situational awareness. In cases where Markers are automatically created by system rules, they are pre-configured with an associated teamwork dimension, based on information defined in the rules that are triggered.

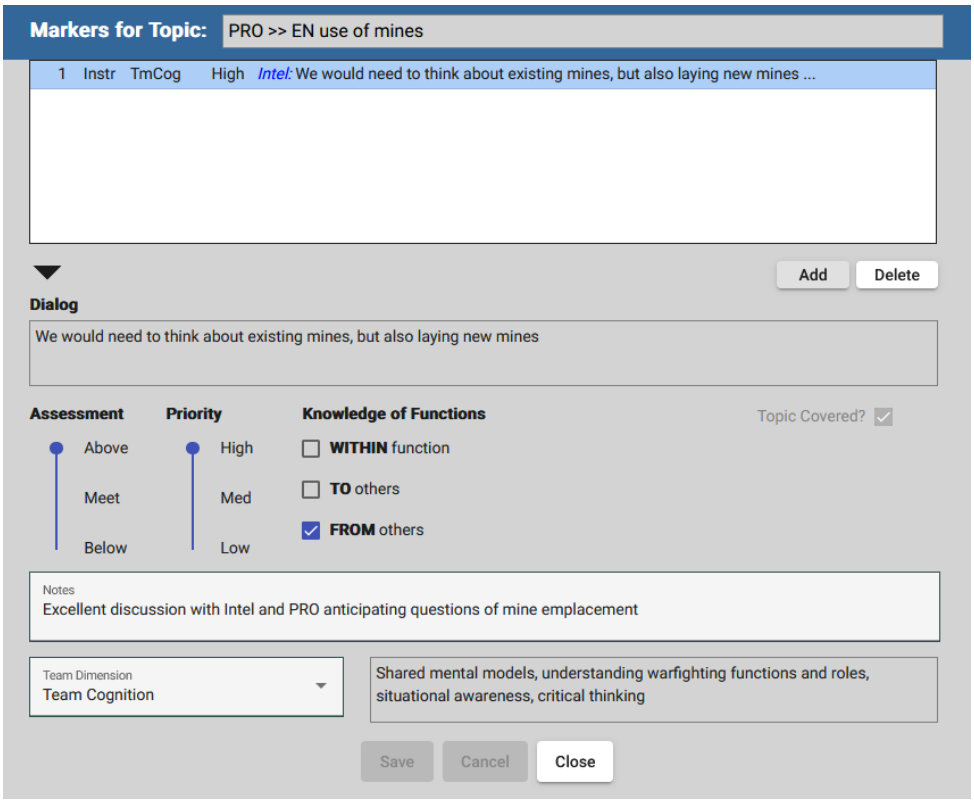

Figure 2. Dialog for Lead / Instructor Review of a Marker

Once the Lead / Instructor determines the exercise is ready to conclude and proceed to AAR, the next step is to notify participants and then go through a process of reviewing Markers to select those to be debriefed. The expected practice is for a subset of Markers will be chosen as AAR Topics, to highlight key examples of good or bad teamwork. During this pause while the Lead / Instructor prepares AAR Topics, participants are encouraged to reflect on their own considerations from the exercise, especially thinking about Knowledge of Functions (WITHIN / TO / FROM their roles and other roles). The AAR is then conducted similarly to the exercise itself, where the Lead / Instructor initiates AAR Topics in the message panel (in a similar fashion to regular exercise Topics) for team discussion. The Lead / Instructor refers to the supplemental information recorded with the Markers that became AAR Topics, to bring out details during this dialog. The review of AAR Topics is then followed by a round-robin where each participant / role discusses their takeaways in general and about Knowledge of Functions.

Implementation

The exercise environment is constructed in the form of a MEAN architecture, which stands for a collection of standard components used to build web applications: MongoDB, Express, AngularJS, and Node.js. One of the implementation objectives is to design for generalizability to other similar domains, and also reusability of data products. To that end, rather than a purely standalone application, the exercise environment is constructed for interoperability with the Generalized Intelligent Framework for Tutors (GIFT, Sottilare et al, 2012), which provides a standards-oriented methodology for organizing learning objectives and data products in training systems, at either the individual or team level. The exercise environment is treated as an external application to GIFT, with communications taking place between the exercise server and a cloud instance of GIFT using REST API calls (representational state transfer application programming interface). GIFT mechanisms are used as the entrypoint for participants to login, receive assigned roles, and join the exercise (Jensen et al., 2023). Training application state messages are then relayed to the GIFT Domain Module during the exercise and at the conclusion. In the initial implementation, the primary data products retained in GIFT are the Markers and also records of Topics covered, along with any supplemental information. On the exercise environment server side, all data from each exercise are retained for future reference and

analytics, although there is no identifying information about participants other than the roles taken and exercise configuration details.

INITIAL EVALUATION

An initial evaluation was conducted with two instructors, to gather subjective input about the training tool and the ways in which it may or may not satisfy training needs, both for the specific application of team familiarization for wargaming preparation, and also in terms of other potential related team training domains. Given the very limited sample, the evaluation was conducted informally, and is not considered a data collection to produce any statistically grounded conclusions. Nonetheless, given the extensive experience of the instructors in this training domain, their observations and feedback are insightful and valuable.

The evaluation was carried out in a single session, and followed up with a survey to elicit feedback in the form of ratings for a range of criteria and written responses to several open questions. The session started with a short introductory sample exercise walkthrough led by the development team, using a scripted set of actions and staff team inputs to illustrate a cross-section of functionality. This was then followed with a fresh exercise session where one of the instructors took the role of the Lead / Instructor, with the other in one of the staff team roles, and members of the development team in other roles. Although the survey is the written record, there was also open discussion during the session. Feedback from instructors covered a range of topics, grouped into the following themes:

- Facilitating the Lead / Instructor during the exercise
- Learning experience for staff team participants
- Data products
- Applications

Each of these themes is discussed in further detail below.

Facilitating the Lead / Instructor

A key takeaway from the instructors' responses is that the ability of the Lead / Instructor to guide the team's discussions and create learning experiences is the most important determinant of training effectiveness. Hence, questions targeted the extent to which the platform supported the Lead / Instructor in performing this function. In general, the instructors agreed that the platform provides a means for the Lead / Instructor to prepare staff for wargaming by familiarizing them with roles and processes, and gaining practice on various scenarios. They expressed the opinion that the platform would allow them to facilitate the exercise and staff participation, and readily support the instructional purpose of tracking staff's responses and actions during the exercise. Instructors also affirmed the importance of having tools to bookmark parts of the exercise to be reviewed as AAR learning points. Relevant comments:

- *"So much of this application's utility will be determined by the Leader / Instructor's ability to draw all the members into a problem-solving scenario so that the team sees the collective power brought by the breadth and depth of their combined thinking skills."*
- *"The application should be fairly intuitive, and easily learned within a few short practice sessions."*

During the evaluation session, there were cases where certain utilities needed explanation, so the second comment above aptly captures the fact that some practice is needed, and also more job aids within the exercise environment would be helpful. There were also several thoughts for refinements or further capabilities. Having used the auto-complete mechanism which draws from a catalog of pre-authored inputs, instructors also saw the potential for the catalog to be augmented with general purpose, frequently used comments that the Lead / Instructor injects to draw out or "pull" information from staff. Since staff inputs during this exercise are often in the form of a question (e.g., "Where would an NAI be placed to detect enemy movements?"), a frequent follow-up from the Lead / Instructor may be an open-ended Prompt to go further (e.g., "What does that mean to our mission success or failure?" or "Why do we need to know?"). Such additions to the auto-complete catalog are easily implemented, and these inputs are very helpful for development, with concrete examples of what instructors would find useful. Instructors similarly expressed interest in further development of support for authoring capabilities ("easy editing rights for the instructor / leader to make adjustments to supporting documents and subtask learning points").

Learning Experience for Participants

This theme relates to how participants learn from the exercise, including questions of whether the learning objectives are clear, how feedback occurs, and what is the role of self-assessment. Instructors agreed that the platform enables staff members to learn how to ask appropriate questions and gain greater understanding of the interdependencies of the various warfighting functions. They concurred that the platform allows participants to collaborate and communicate naturally, and exhibit different teamwork behaviors. Relevant comments:

- *"This application allows for the team to discuss how to approach elements of problem identification, analysis of problem elements, key variables, and requirement to solve the elements."*
- *"This is suited for both resident and distant learning. However, optimally this is best suited for distant learners."*

Since the processes involved in wargaming and COA analysis are high on task and outcome interdependence, one of the challenges in constructing a team exercise is to clearly convey the learning objectives that relate to team dimensions, which can be somewhat abstract. As a general principle, instructors emphasize the need to make learning objectives clear to the training audience, along with how they will be assessed. A command staff lead with a background in Army aviation may be naturally oriented to think of aviation-related tactical considerations and learning objectives, much more than team cognition or supporting behaviors. In the current implementation, although the exercise provides opportunities to exhibit teamwork in an applied scenario, there is little to prime the exercise for participants to think about teamwork behaviors as the learning objectives. This is an area for planned improvement.

Instructors emphasized the value of reflection and self-evaluation in learning. Although they agreed that the platform would support learning in different ways, they suggested the introduction of a more explicit mechanism for participants to complete and share post-exercise self-assessments, and specifically articulate include new discoveries made about their own functional role and others. This is part of the intended use under the facilitation of the Lead / Instructor, but implementation to date does not include a specific mechanism like this, so this is another valuable input for further development. Instructors suggested that a list of prompts and responses to participants by role at end of exercise would add to their understanding of the various warfighting functions.

Data Products

Questions about data products explored the types of data and records from the platform that would be of interest to various stakeholders. These include the Lead / Instructor, the student, curriculum developers, C2 training system engineers, TRADOC stakeholders, as well as researchers in the training community. The instructors' responses centered around the interests of the Lead / Instructor and student, such as the log of Prompts and responses, the in-exercise assessments made by the Leader with Markers, and potential post-exercise self-assessments by the student. These were to aid the Lead / Instructor with AAR and training, and to provide students with opportunities to affirm their own learning and learn from each other. Relevant comments:

- *"A copy of the prompted questions and the responses by role (function) may be of value when students rotate positions and or have to fulfill a role they are not familiar with."*
- *"Perhaps the most valuable record for team members would be sharing of post-instruction self-assessments made by each member."*

The instructors show little interest in data products associated with quantitative outcome scores to grade teamwork. Based on their initial feedback, more of the focus is on the process and gains from the exercise itself and the opportunity for team self-assessment, rather than preserving numerical scores. So although there may be some metrics used to quantify exercise data such as Markers and Topic coverage, instructors spoke mainly in terms of the knowledge products (e.g., "sharing of post-instruction self-assessments" above). Accumulated exercise data may also be of value to other stakeholders. For instance, data by exercise scenario may show which critical events in the COA tend to require collaboration from which warfighting functions, and how. Semantic analysis of the prompts and messages by student experience levels may lead to insights into how tactics and strategies develop, as students become more competent. Data products at a more granular level may also be useful to support various analytics, like how often the instructors modify the pre-authored Prompts before sending to students, how long messages tend to be, how much time tends to be devoted to each Topic, and so on. Instructors also considered future uses for data to help refine the

exercise itself over time. The message logs from exercises conducted over time contain free-text examples that can be automatically mined to expand the catalog of auto-complete messages available to the Lead / Instructor and other roles. The accompanying data from Markers helps to associate qualitative judgements with messages, for a machine learning approach to expanding the catalog. This is not implemented, but another valuable future consideration.

Applications

Questions about the exercise's applicability were initially framed in terms of training command staff at battalion and higher echelons specifically at the Command and General Staff College (CGSC) at Ft. Leavenworth. However, discussion expanded beyond this scope in terms of applicability to other team collaborative problem-solving domains, and also into more general concepts of what works for such domains without specifically limiting the discussion to this particular implementation of this wargaming preparation exercise. For the immediate applications, a relevant comment was that the training system would be helpful for:

- *"Captains Career Course, and CGSC level staff training. Also helpful for Reserve Component staff training because many of their subunits are in geographically dispersed locations."*

As mentioned above, the instructors' feedback about applicability expanded beyond the context of this training tool, to more of a discussion of how *this kind* of exercise can be productive in a variety of team problem-solving tasks. In the same way that the wargaming preparation and team familiarization exercise is a simplification of a full-scale wargaming exercise, other team problem-solving domains would likely benefit from a smaller scale exercise to work on understanding cross-functional roles as preparation for the full scope of a larger exercise or operation. Some of the common features or desired functionality for exercises to meet this purpose include:

- Participants to access background information and resource materials
- Participants to view and refer to a common display or representation of the problem space
- Participants to be aware of the status of other participants
- Participants have various ways to communicate and for their discussions to be tracked and threaded
- If conducted using chat or messaging utilities, participants have the option to use domain-specific auto-completes, which saves time and can be used to provide cues or hints
- Participants have options to display different levels of engagement throughout the discussion
- The leader or facilitator to note observations and assessments during the task
- The leader or facilitator to use available pre-populated discussion points, with flexibility to adapt sequencing and content

Domains for which such a platform could be used include cross-functional taskforces, cross-functional business teams, engineering teams, product design teams, emergency response teams, and a variety of virtual teams. Instructors also commented that exercises oriented toward cross-functional team familiarization have applications outside of training. Citing examples with Army operational units, especially where command staff teams may composed on the fly, a small-scale distributed preparatory exercise may help develop patterns of collaboration before entering into operations.

ANALYSIS AND CONCLUSION

The road ahead includes future plans to conduct pilot studies with students. This will be essential to evaluating training benefits, especially for the direct purpose of preparing staff members for wargaming. Since there is an elective course at CGSC focused on wargaming preparation, this is likely an ideal venue for a pilot study. In the near term for continued development, some of the most direct takeaways from the instructors' initial evaluation feedback involve immediate next steps in further developing or refining the wargaming preparation training tool. But also much of the feedback is motivated by more general purpose observations that may apply for other similar applications. Aside from the mostly positive initial feedback, the following is a review of some of the suggestions for improvement:

- **Instructor authoring tools**. Ideally a deployed training tool supports instructor authoring, so that new scenarios and customizations can be easily added without relying on a contractor. This is a natural future step for development, but very limited in the current version of the tool.

- **Expanded instructor Prompts**. Although scenarios are developed with pre-authored Prompts that instructors use to initiate Topics, there are also general purpose, commonly used follow-up Prompts that would be helpful to have available in the auto-complete catalog for the Lead / Instructor.
- **Leveraged exercise data for auto-completes**. Data-driven methods can be used to automatically expand the auto-complete catalog over time, ingesting examples of free-text inputs from exercise data logs.
- **Self-assessment utilities**. With the instructors' emphasis on the importance of team self-assessments, there is a need to integrate specific tools to elicit these self-assessments from exercise participants, structured in terms of thinking about gained knowledge of functions (within / to / from functional areas).
- **Data analytics**. Planning for a future deployed version of the training tool should include the development of analytics utilities that instructors and other stakeholders can use to make different kinds of measurements of trends or statistics from accumulated exercise data.

Another area for further work is in multilevel assessments. Although the wargaming preparation training tool is developed to facilitate team training and many assessments are at the team level, individual assessments will also be relevant. This is because teams can only improve if individual members improve. To obtain individual-level assessments from team-level assessments, we need to understand the nature of task interdependence and patterns of interactions among the team during task performance (Cannon-Bowers & Bowers, 2011; Kozlowski & Chao, 2012; Kozlowski & Klein, 2000). Chen, Mathieu, and Bliese (2005) propose several models describing possible interactions and dynamics within the team that warrant different ways to derive multilevel assessments. For instance, the individual-level score should not simply be an equal distribution of the team score if certain members played a more dominant role in the team task (e.g., a "superstar" staff member who drives the entire discussion). In such a case, the team score may be more reflective of the most dominant member than of all members. The model for such a competency construct would be the select score model, since a key member determines most of the team's performance. Other models are the aggregate, summary index, consensus, referent-shift, and dispersion models (Chan, 1998; Chen et al., 2005). Comprehensive analysis of team interactions, possibly through social network analysis, can provide a basis for identifying the most appropriate model for the construct of interest.

Regarding the browser-based platform for the training tool, instructors weighed in on the tradeoff between enabling distributed training and selecting chat-style messaging as the primary monitored form of communication. Although the instructors acknowledge the limitations imposed on natural team interaction, they also place significant value on the ability to support team training events with remote participants – both geographically remote and also in the nearby-but-off-hours context of homework assignments outside the classroom.

One of the other findings gleaned from the feedback given by instructors is the suggestion that a smaller-scale preparatory team familiarization exercise may be an advisable option for many domains, both training and operational. Whereas the development attention for complex domains is often focused on creating complex exercises that can mirror the level of detail in the real-world, a small-scale exercise can have the advantage of focusing on a specific gap such as cross-functional teamwork, with reduced exercise overhead and more opportunities for practice. This observation may be a consideration for programs creating new infrastructure and tools for a wide range of training applications, such as the GIFT framework and the Army's Synthetic Training Environment (STE).

Finally, for the team of researchers working on the concepts, design, and implementation of the wargaming preparation exercise described in this paper, there is a parallel between studying it and living it. In a practical sense, the development team is its own example of a cross-functional collection of participants contributing different areas of expertise for a complex collaborative problem-solving challenge. The research team includes a mix of cognitive psychologists versed in human factors and learning science, developers with knowledge of different technical disciplines, subject matter experts with years of applied instructional and operational experience, and project leads to help coordinate across roles. It is an important challenge to build tools to help command staff teams develop teamwork skills they will ultimately use to plan and execute optimal COAs that save resources and lives on the battlefield. This is a motivator for the research team, who have the same need to engage in effective cross-functional processes that involve team dimensional behaviors, similar to those to be fostered in the training audience.

ACKNOWLEDGEMENTS

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 representing official positions, either expressed or implied, of the U.S. Government.

REFERENCES

Cannon-Bowers, J. A., Tannenbaum, S. I., Salas, E., & Volpe, C. E. (1995). Defining team competencies and establishing team training requirements. In R. Guzzo & E. Salas (Eds.), *Team effectiveness and decision making in organizations* (pp. 333-380). San Francisco: Jossey-Bass.

Cannon-Bowers, J. A., & Bowers, C. (2011). Team development and functioning. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology, Vol. 1. Building and developing the organization* (pp. 597–650). American Psychological Association. https://doi.org/10.1037/12169-019

Chan, D. (1998). Functional Relations Among Constructs in the Same Content Domain at Different Levels of Analysis: A Typology of Composition Models. *Journal of Applied Psychology*, *83*(2), 234–246.

Chen, G., Mathieu, J. E., & Bliese, P. D. (2005). A framework for conducting multi-level construct validation. In *Multi-level issues in organizational behavior and processes* (pp. 273–303). Emerald Group Publishing Limited.

Cianciolo, A. T., & Sanders, W. R. (2006). *Wargaming Effectiveness: Its Conceptualization and Assessment* (Technical Report No. 1178; p. 128). U.S. Army Research Institute for the Behavioral and Social Sciences.

Cohen, M. S., & Thompson, B. B. (2001). 6. Training teams to take initiative: Critical thinking in novel situations. In *Advances in Human Performance and Cognitive Engineering Research* (Vol. 1, pp. 251–291). Emerald (MCB UP). https://doi.org/10.1016/S1479-3601(01)01008-6

Cooke, N. J., Gorman, J. C.,&Rowe, L. J. (2008). An ecological perspective on team cognition. In E. Salas, G. F. Goodwin & C. S. Burke (Eds.), *Team effectiveness in complex organizations: Cross-disciplinary perspectives and approaches*, 157–182. New York: Routledge/Taylor & Francis Group.

Courtright, S. H., Thurgood, G. R., Stewart, G. L., & Pierotti, A. J. (2015). Structural interdependence in teams: An integrative framework and meta-analysis. *Journal of Applied Psychology*, *100*(6), 1825–1846. https://doi.org/10.1037/apl0000027

Fischer, S. C., Spiker, V. A., Harris, D. H., McPeters, E. R., & Riedel, S. L. (2008). *Computerized training in critical thinking (CT) 2: A skill-based program for Army personnel*. ANACAPA SCIENCES INC SANTA BARBARA CA.

Gibson, C. B., & Cohen, S. G. (Eds.). (2003). *Virtual teams that work: Creating conditions for virtual team effectiveness* (1st ed). Jossey-Bass.

Headquarters, Department of the Army. (2017). *Operations* (ADP 3-0; Army Doctrine Publication).

Jensen, R., Finnigan, E., Teo, G. & Goodwin, G. (2022) Team Assessment Methods for Training Command Staff Decision Making. *Proceedings of the 10th Annual GIFT Users Symposium (GIFTSym10)*.

Jensen, R., Finnigan, E., Teo, G. & Townsend, L. (2023) Refining a Team Training Approach to Prepare Command Staff for Wargaming. *Proceedings of the 11th Annual GIFT Users Symposium (GIFTSym11)*.

Kozlowski, S. W., & Chao, G. T. (2012). The dynamics of emergence: Cognition and cohesion in work teams. *Managerial and Decision Economics*, *33*(5–6), 335–354.

Kozlowski, S. W., & Klein, K. J. (2000). *A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes.*

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *The Journal of Applied Psychology*, *85*(2), 273–283. https://doi.org/10.1037/0021-9010.85.2.273

Shuffler, M. L., Diazgranados, D., Maynard, M. T., & Salas, E. (2018). Developing, Sustaining, and Maximizing Team Effectiveness: An Integrative, Dynamic Perspective of Team Development Interventions. Academy of Management Annals 2018, Vol. 12, No. 2, 1–37. https://doi.org/10.5465/annals.2016.0045

Sottilare, R., Brawner, K., Goldberg, B. & Holden, H. (2012). The Generalized Intelligent Framework for Tutoring (GIFT). US Army Research Laboratory.

Tannenbaum, S. I., & Cerasoli, C. P. (2013). Do Team and Individual Debriefs Enhance Performance? A Meta-Analysis. *Human Factors*, *55*(1), 231–245. https://doi.org/10.1177/0018720812448394

Teo, G., Jensen, R., Johnston, J., DeFalco, J., & Goodwin, G. (2021). Measures for Assessing Command Staff Team Performance in Wargaming Training. *Proceedings of the Interservice / Industry Training, Simulation, and Education Conference (I/ITSEC 2021)*.