Designing Assessments in a Team Trainer for Wargaming

Grace Teo[1], Randy Jensen[2], and Gregory Goodwin[3]

[1]Quantum Improvements Consulting, Orlando, USA
gteo@quantumimprovements.net
[2]Stottler Henke Associates, Inc., San Mateo, USA
jensen@shai.com
[3]U.S. Army Combat Capabilities Development Command – Soldier Center, Orlando, USA
gregory.a.goodwin6.civ@army.mil

Abstract. This paper discusses several concepts for the development of a distributed trainer for command staff trainees learning to develop courses of action (COAs) and wargame. These concepts include how understanding the nature of the team tasks determines the taskwork and teamwork competencies and shapes the pedagogical strategies to be incorporated into the trainer. As well as concepts related to the difficulties in developing assessments for unstructured team tasks and the challenges with assessing team processes, we also discuss the inclination towards a positivist paradigm that relies on the presence of behaviors for indicators, when absence of certain behaviors can also be indicative and used in assessments. We conclude with a preliminary framework for organizing system features for the trainer, and ideas for future research.

Keywords: Taskwork, Teamwork, Pedagogical Strategies, Assessments

1 Introduction

Wargaming is a team activity conducted by command staff with expertise in various warfighting functions (WfFs). The team develops and analyzes courses of action (COAs) by considering critical events, actions, reactions, and counteractions. Their analysis typically results in an operational plan. Any instructional system developed for training wargaming must enable team task execution while also serving training objectives. The extent to which these two goals are met depends on the design of the assessments as well as the training experience and flow of the training exercise delivered by system features. In this paper, we share progress in the design and development of a prototype distributed trainer for U.S. Army wargaming. Our approach to designing assessments for the prototype involves understanding the task and competencies to be trained, and defining indicators or markers of those competencies for assessments. These can only be implemented with system features that can elicit indicative behaviors during task execution and training.

2 The nature of the wargaming team task

The type of team task to be trained should drive many of the system requirements for any trainer. Identifying the type of team task can help researchers and system developers gauge the typical flow of information and communications among team members and anticipate patterns of interactions so that requirements can be designed accordingly. A typology of team tasks had been proposed [1] which describes the ways in which team members work together to accomplish the task. The typology includes characteristics such as complexity, routineness, type of interdependence members have on each other which can be *pooled*, *sequential*, *reciprocal*, or *team* [2]. For instance, members of a team working the assembly line show *sequential* interdependence, while co-authors on a writing team exhibit *reciprocal* interdependence as the writing is passed back and forth between members. Team tasks can also vary in whether they focus on managing, advising, negotiating, performing a service, executing a psychomotor action, or solving problems which can be defined or ill-defined [3]. The wargaming team task can be considered a complex, ill-defined problem-solving task characterized by *team* interdependence where members' tasks and work sequences are unspecified and dynamic. Although it is a challenge to derive any interaction structure for such an unstructured team task with no defined stages requiring a high level of interdependency, there is semblance of some turn-taking in the wargaming team task. For instance, in Division level wargaming, the Chief of Staff typically directs the discussion of COA events and phases, and the staff representing the Intel WfF tends to be called on first to provide background information with intelligence about the enemy. From there, other staff members perform their specialized WfF roles by contributing critical function-specific information as the team steps through various possibilities and topics in their COA discussion. Each team member must not only be familiar with how the different WfFs work together both in planning and execution, but have a nuanced understanding of the COA scenario, including possible repercussions of hypothesized events, and how different WfFs impact further COA decisions down the line.

3 Competencies and pedagogical strategies

An application that enables the command staff team to prepare for and conduct wargaming may not sufficiently support training if it does not support the acquisition of the skills and competencies needed for the task. The team's training must include taskwork and teamwork, and incorporate assessments of outcomes and processes.

3.1 Taskwork and teamwork

While taskwork pertains to *what* the team members do to achieve the collective goal, teamwork focuses on *how* they interact and collaborate to accomplish the team task. A trainer for wargaming should facilitate both taskwork and teamwork in a way that is appropriate for each team member according to their role [4]. Teams that consistently develop good COAs and subsequently wargame effectively would have mastered the needed taskwork and teamwork competencies. There are different levels of taskwork that the trainer should assess; scenario-specific responses such as selecting a "route along the coast which bypasses a mountainous range", and concepts or principles that

drive the scenario-specific response, which include assumptions and presuppositions held knowingly or unknowingly by the trainee. These principles are more difficult to assess since they are rarely evident from superficial responses, and may require extra probing before they are apparent. For instance, a further prompt such as "what is the rationale for this selection?" may reveal that the coastal route was selected mainly for speed rather than to minimize risk. If a safer but equally fast route was available, then the coastal route would have been suboptimal for meeting the stated purpose. A trainee may arrive at a suitable scenario-specific response as a mistaken application of a rationale or principle. In wargaming training, it is more crucial to uncover these errors in decision rationale rather than to look for a nominally correct scenario-specific answer, especially for tactical decisions where many options may be acceptable. As these principles are scenario-agnostic, there is a possibility of developing such prompts which can be applied across different scenarios. Instructors and Subject Matter Experts (SMEs) in wargaming training have developed a "mental model matrix" to capture some of these scenario-agnostic principles and topics to address.

The matrix lays out the critical topical areas staff teams must consider for any COA, with a breakdown by WfF. Whether and how trainees address these topics may reveal certain patterns of thinking, including cognitive biases. For example, named areas of interest (NAIs) are intelligence collection points associated with specific locations, typically planned by an Intelligence lead. In explaining how NAI placements contribute to maintaining contact with the enemy, an Intel staff member may realize that they tend to seek out confirmatory evidence. An exercise that involves prompting trainees in this manner can help provide the training experience that promotes development of the right mental models and metacognitive processes.

While instructors and SMEs seek to develop such critical thinking skills and metacognitive awareness in their trainees, they also recognize that the group dynamics within the team can substantially impact how the knowledge and skills of individual WfF staff roles are manifested, drawn out, and sharpened within the team. These interactions constitute the teamwork aspect or process by which the team performs the task. Wargaming assessments are often based on outcome measures such as the synchronization matrix showing the final mission plan and details of dependencies, or the quality of the COAs developed and selected based on criteria such as feasibility, suitability, and completeness. While these may reflect the quality of the team's process, they are not direct measures of it. Training assessments can only provide insight into the team's teamwork if the trainer captures indicators of the team's process, which in turn requires the team competencies to first be defined.

In our approach to ensure that the trainer supports the teamwork needed in wargaming, we identified the team competencies most relevant to COA development and wargaming analysis. These were leadership, team cognition, information exchange, communication quality, supporting behaviors, and team orientation. Teamwork during wargaming involves guiding and directing (leadership), cooperating and offering support (supporting behaviors), and sharing information with each other (information exchange) that builds a shared mental model (team cognition) in a clear and appropriate manner (communication quality) which shows trust and openness (team orientation).

In training for these taskwork and teamwork competencies, instructors and SMEs emphasize the importance of having trainees master the roles of the various WfFs and understand how these must work closely together. They typically adopt the "Socratic"

method where trainees are guided through shared dialogs to pose questions, evaluate the reliability of incoming information, cross-examine assumptions by generating alternative explanations or seeking disconfirming evidence [5]. This type of pedagogical approach does not necessarily require high fidelity simulations or even lengthy and complex scenarios. It can be supported by a relatively open system architecture that includes a clear depiction of the vignette of interest, and allows members to respond to open-ended prompts, raise objections and questions, work together to identify decision points, and learn from and build on each other's contributions. From our analysis of the tasks and of instructional methods, we propose that potential pedagogical strategies that support the acquisition of the taskwork and teamwork competencies for wargaming include the following (see Table 1):

Table 1. Summary of wargaming competencies and suggested pedagogical strategy

	Competencies	**Pedagogical/instructional strategy for acquiring competencies**
Taskwork	Individual's knowledge of WfFs, the wargaming process and the military decision making process (MDMP), military protocols and conventions, critical thinking skills, ability to perform assigned staff role	Application of various knowledge and skills to a wide range of scenarios/vignettes and echelons.
Teamwork	Leadership in the team, team cognition, information exchange, team orientation, supporting behaviors, communication quality	Group discussions and questions that reveal and clarify preconceptions, generate ideas to test hypotheses, identify decision points. Active listening practice. Cross-training on staff roles.

It is possible that novice command teams would exhibit different patterns of interactions from expert teams, as will teams composed of members from different services. For instance, all-Army, all-Navy, or mixed-services which will be prevalent in multi-domain operations (MDO), may show different patterns of collaboration due to differences in their tactics, techniques, and procedures (TTPs). Artificial intelligence (AI) can help extract features in team interactions that characterize how different types of teams wargame various types of vignettes and scenarios such as MDOs, operations across echelons, etc.

4 Pre-defining indicators for assessments

In designing assessments for our trainer prototype, we drew from concepts in the Event-Based Approach to Training methodology (EBAT)[6, 7], and the Event Analysis of Systemic Teamwork methodology (EAST)[8, 9]. These methodologies have been successfully applied to training command and control teamwork in aviation and military domains [10–13]. The EBAT involves systematically identifying and injecting events in the training exercise to elicit pre-defined opportunities for observing behaviors indicative of constructs of interest and training objectives. This encourages traceability from behavioral indicators to assessments, and training objectives [14]. The EAST methodology proposes making explicit (i) who the members in the exercise are,

(ii) when tasks are performed and who they involve, (iii) where members are, (iv) how members collaborate and communicate to achieve task goals, and (v) what tasks members are performing, and what knowledge and information is shared and used [12]. Both the EBAT and EAST advocate explicitly for articulating and pre-defining anticipated behaviors and markers that serve as measures for constructs of interest. Whereas an event in the EBAT and EAST in an example aviation task may be "reaching cruising altitude" or "initiating the landing procedure", events in dialog-based, collaborative wargaming preparation tasks could be the prompted discussion of a COA decision point such as anticipating the impact of enemy reinforcements on the scheme of maneuver. Assessment measures for this kind of event concern how well the team discussion covered relevant tactical, cross-functional considerations.

To some extent, such an approach implies a positivist research paradigm which emphasizes positive observations or the presence of behavioral evidence, although the absence of behaviors can also be indicative and should be included in assessments. However, this requires specifying a priori expectations, which are challenging for tasks that are unstructured. For instance, in a well-defined and structured task such as a maintenance task, we can assess absence of certain desirable behaviors when they are not observed in the procedural steps. In unstructured tasks, there are fewer expectations of when certain behaviors should be exhibited, so it is more difficult to note when these are absent. Table 2 presents examples of "positive" (presence of behaviors) and "negative" (absence of behaviors) observations that can be indicators of the teamwork dimensions identified previously. Some are contextualized to wargaming training in a primarily dialog-based exercise environment, and some are more general purpose.

Table 2. Examples of indicators from teams high and low on each team dimension

Team Dimension	Presence of these behaviors	Absence of these behaviors
Leadership (guidance, direction, coordination, strategy formulation)	*High on dimension* -Leader guides who should be doing what, and when. Active team member participation. *Low on dimension* -Frequent questions on where team is at within the exercise, staff looking at the wrong information	*High on dimension* -Absence of team behaviors indicating boredom or distraction *Low on dimension* -Absence of leader probing questions or indicators of active team engagement
Team cognition (knowing who knows/needs what and the WfF roles, critical thinking, metacognitive awareness)	*High on dimension* -Addresses the right role for information and questions -Discussion appropriate for the echelon -Shows understanding of interdependencies among WfFs, (e.g., "Signal's input is needed about maintaining comms, if Aviation takes this helicopter route") -Questions assumptions, generates alternative explanations, detects missing information, seeks disconfirming evidence	*High on dimension* -Infrequent requests for clarification on staff roles

Team Dimension	Presence of these behaviors	Absence of these behaviors
Team cognition *(continued)*	*Low on dimension* -Addresses the wrong WfF role for questions and information -Jumps to conclusions, reaches for easy explanations, too focused on a quick resolution even with new information	*Low on dimension* -Members not seeking disconfirming evidence, not seeking or using information from other roles
Information Exchange (knowing what info. is needed, how much detail is needed, when to give it)	*High on dimension* -Volunteers information to the right role in a timely manner, with enough detail for the echelon -Discusses topics relevant to the echelon *Low on dimension* -Gives incomplete, untimely, or inaccurate information -Gives information for the wrong echelon	*High on dimension* -Absence of excessive requests for information ("pulling information") from staff *Low on dimension* -Absence of appropriate questions, not discussing relevant topics
Supporting Behaviors (backup behaviors, load-leveling, mutual performance monitoring, giving feedback)	*High on dimension* -Shares information that assists others in their work or role -Reminds team of important information missed or overlooked *Low on dimension* -Adds to others' work even when they are busy -Asks unnecessary questions that distract from topic at hand	*High on dimension* -Absence of distracting questions or comments *Low on dimension* -Failure to assist or ignores others when they need help
Team Orientation (promotes and supports open communication that facilitates mutual trust, team cohesion, team motivation, conflict resolution)	*High on dimension* -Responsive to each other, shows support (e.g., *like* button) -Encourages contributions, uses "we" often, shares credit/blame as a team -De-escalates and resolves conflicts *Low on dimension* -Quick to claim credit, shifts blame -Defensive when questioned, pushes own ideas.	*High on dimension* -Not dismissive of others -Does not fuel conflicts *Low on dimension* -Fails to attend to or acknowledge others when they contribute -Lack of participation
Communication Quality (use of proper phraseology, awareness of military conventions)	*High on dimension* -Uses standard conventions and protocols that facilitate clear communications -Communicates concisely -Adjusts communication style as needed *Low on dimension* -Uses wrong terminology that may cause confusion -Uses wrong channels, gives information in a roundabout way. -Engages in unnecessary chatter	*High on dimension* -Absence of unnecessary chatter *Low on dimension* -Fails to acknowledge others, no closed-loop communications

These measures can be further expanded with more data and research with AI and machine learning (ML). For instance, AI-based methods may detect changes in the frequency and type of questions discussed by a team as they coalesce and work better together. Novel assessment measures may be derived from applying AI and ML approaches to extract emotion indicators from facial expressions, gestures, eye tracking, or vocal data. AI-based speech recognition methods can potentially assist in automating assessments as well.

5 System features for a wargaming trainer

It is relatively straightforward to assess the team's taskwork from responses and outcome measures such as ratings of the tactical decisions within COAs based on criteria set out in rubrics. To assess a team's processes and teamwork, the wargaming trainer must collect data of the members' interactions. These interactions can be scripted into the workflow or be unscripted and ad hoc. Scripted interactions can be embedded in the workflow of the exercise if the team members are required to respond to other members' inputs. These interactions offer the best opportunities for automated assessments. Ad hoc interactions are extemporaneous and can be initiated by any member at any time throughout the team activity. The exercise flow should be designed to create opportunities for both types of interactions to elicit a wide range of behaviors, some of which are indicators of the team competencies. Although the technology in the trainer can be leveraged to automate assessments, assessment opportunities for ad hoc interactions require observer-based assessments. Regardless of how much the interactions are scripted, they are initiated by members receiving information and involve them responding and taking action after processing the stimuli. Given this inevitability, we propose the following framework for system features for the team trainer (see Table 3). Such a framework can serve as a guideline for system development.

Table 3. System features for wargaming taskwork and teamwork

	Features to deliver stimuli	**Features that accept inputs**
Taskwork	-Display to allow team members to obtain information needed for their roles. *E.g., Intel staff needs information on Commander's critical information requirements (CCIR), and priority intelligence requirements (PIR)* -List of topics to cover in discussion and prompts. *E.g., "What do you need to consider for this situation and what roles are involved?"*	-Interface for submitting products of taskwork, or for assigning taskwork. *E.g., drop-down menu of possible responses and auto-complete options, text box for free form text*
Teamwork	-Display to promote shared understanding across roles. *E.g., common map of area of interest (AOI)* -Display needs to promote awareness of others' status. *E.g., status board showing current status of all staff members*	-Interface to allow initiation of action directed at other members or responding. *E.g., drop-down menu of possible responses and auto-complete options, text box for free form text, messaging with select member(s), "like" button to endorse or acknowledge*

6 Future work and conclusion

Building a trainer to prepare command staff for collaborative wargaming requires attention to both task execution and achieving learning objectives. As with most team tasks, there are taskwork and teamwork competencies to be developed in wargaming training. The challenge lies in the fact that this team task involves few tangible artifacts and does not readily adhere to any type of structure. Being in a cognitive domain of learning [14, 15], the training focus is on the abstract and conceptual, and is difficult to operationalize and assess. In addressing the concepts discussed in the paper, we identified opportunities to apply AI in wargaming training that include learning how types of teams (e.g., novice, expert, mixed services teams, all-Army teams) interact and communicate, exploring if there are patterns of interactions that can be extracted for certain types of vignettes (e.g., operations for different echelons, multi-domain operations), training speech-recognition for the wargaming domain, and developing assessments from novel measures such as eyetracking and gesture and facial expression recognition. We hope to use the prototype currently under development to collect data that can help refine some of these research opportunities.

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.

References

1. Bonner, D., Gilbert, S., Dorneich, M. C., Burke, S., Walton, J., Ray, C., & Winer, E. (2015). Taxonomy of Teams, Team Tasks, and Tutors, 9.
2. Saavedra, R., Earley, P. C., & Van Dyne, L. (1993). Complex interdependence in task-performing groups. *Journal of applied psychology*, *78*(1), 61.
3. Wildman, J. L., Thayer, A. L., Rosen, M. A., Salas, E., Mathieu, J. E., & Rayne, S. R. (2012). Task types and team-level attributes: Synthesis of team classification literature. *Human Resource Development Review*, *11*(1), 97–129.
4. Crawford, E. R., & Lepine, J. A. (2013). A Configural Theory of Team Processes: Accounting for the Structure of Taskwork and Teamwork. *Academy of Management Review*, *38*(1), 32–48. https://doi.org/10.5465/amr.2011.0206
5. Delic, H., & Bećirović, S. (2016). Socratic Method as an Approach to Teaching. *European Researcher*, *111*, 511–517. https://doi.org/10.13187/er.2016.111.511
6. Fowlkes, J., Dwyer, D. J., Oser, R. L., & Salas, E. (1998). Event-Based Approach to Training (EBAT). *The International Journal of Aviation Psychology*, *8*(3), 209–221.
7. Rosen, M. A., Salas, E., Wu, T. S., Silvestri, S., Lazzara, E. H., Lyons, R., … King, H. B. (2008). Promoting teamwork: An event-based approach to simulation-based teamwork training for emergency medicine residents. *Academic Emergency Medicine*, *15*(11), 1190–1198. https://doi.org/10.1111/j.1553-2712.2008.00180.x
8. Stanton, N., Baber, C., & Harris, D. (2008). *Modelling Command and Control: Event Analysis of Systemic Teamwork*. Ashgate Publishing, Ltd.
9. Walker, G. H., Gibson, H., Stanton, N. A., Baber, C., Salmon, P., & Green, D. (2006). Event analysis of systemic teamwork (EAST): a novel integration of ergonomics methods to analyse C4i activity. *Ergonomics*, *49*(12–13), 1345–1369.
10. Fowlkes, J. E., Lane, N. E., Salas, E., Franz, T., & Oser, R. (1984). Improving the Measurement of Team Performance: The TARGETs Methodology, 16.
11. Harris, D., & Stanton, N. A. (2010). Aviation as a system of systems: Preface to the special issue of

human factors in aviation. Taylor & Francis.
12. Salmon, P. M., Lenne, M. G., Walker, G. H., Stanton, N. A., & Filtness, A. (2014). Using the Event Analysis of Systemic Teamwork (EAST) to explore conflicts between different road user groups when making right hand turns at urban intersections. *Ergonomics*, *57*(11), 1628–1642.
13. Smith-Jentsch, K. A., Zeisig, R. L., Acton, B., & McPherson, J. A. (1998). Team dimensional training: A strategy for guided team self-correction.
14. Gagné, R. M. (1972). Domains of learning. *Interchange*, *3*(1), 1–8.
15. Hoque, M. E. (2016). Three domains of learning: Cognitive, affective and psychomotor. *The Journal of EFL Education and Research*, *2*(2), 45–52.