Teamwork Training Architecture, Scenarios, and Measures in GIFT

Robert McCormack1, Tara Kilcullen1, Anne M.Sinatra2, Alexander Case1, Daniel Howard1

Aptima, Inc.1, U.S. Army Combat Capabilities Development Command (CCDC)- Solider Center - Simulation & Training Technology Center (STTC)2

INTRODUCTION

The success or failure of teams facing complex tasks often hinges on their ability to communicate and collaborate throughout the mission. Indeed, numerous studies have shown that developing and maintaining strong teamwork skills through training interventions have a positive impact on performance outcomes (Sottilare, et. al. 2017; Wilson, et. al. 2007; Salas, et. al. 2008). The U.S. Army has specifically recognized the impact of teamwork on mission command and the need for commanders to build teams rather than relying on pre-established relationships. Furthermore, effective teams are those that can establish a high degree of coordination both within their unit and across higher, lower, and adjacent echelons (ARDP 6-0, 2012). Training teams to improve their teamwork skills, such as coordination, however, is often treated as an implicit by-product of task-focused training rather than a goal in and of itself. As such, the ability for trainers to explicitly measure and intervene in teamwork skill development is often lacking. Intelligent Tutoring System (ITS) frameworks, such as the Generalized Intelligent Framework for Tutoring (GIFT; Sottilare et. al., 2012; Sottilare et. al., 2017), offer a great deal of potential for team training, but have generally been developed with individual training in mind. In a previous GIFTSym paper (McCormack, et. al., 2018), the authors presented a set of proposed scenarios and an approach to developing teamwork measures within a virtual training environment. This paper builds upon that work and describes an effort to extend GIFT to include the necessary components of teamwork training. We discuss a GIFT architecture that provides team-level skill feedback, while minimizing the infrastructure cost of developing and main- taining Domain Knowledge Files (DKFs). In addition, we present a number of realistic, doctrinally-rele- vant scenario vignettes developed to provide multiple opportunities for teamwork skill development and feedback. Finally, we provide examples of measures of coordination tailored to the scenario that enable analysis and feedback on skill development.

While the ultimate goal is to enable GIFT to support teamwork training across multiple virtual and live environments, the initial focus of this effort was on training Army teams within Virtual Battlespace (VBS) 3.0. This virtual training environment (VTE) offers multiple advantages: it provides a realistic backdrop for doctrinally-relevant tasks; it enables scalable team sizes and multiplayer experiences; and, much of the infrastructure for enabling integration of VBS and GIFT has already been developed.

Team Training Architecture

In order to instantiate teamwork measures in GIFT, we developed an architecture that would enable meas- urement across multiple VBS players, without inducing a heavy burden associated with creating and man- aging multiple DKFs. Previous efforts at teamwork measurement required a DKF for each player and ad- ditional DKFs for each combination of players and the team as whole (Bonner, et.al., 2017). Any measure- ments for combinations of dyads, triads, etc. of individuals required a separate DKF. While this allows player and sub-team specific measurement, as the team size grows, the exponential growth of DKFs re- quired by this approach quickly becomes unmanageable. Because our focus is on team-level measures, we chose an architecture that was best suited for that, but still allows future inclusion of individual measures.



Figure 1. Team Training Architecture and DKF use

[Figure 1](#_bookmark0) provides an overview of the team training architecture. Each team member runs a local GIFT instance as well as VBS. One team member serves as the VBS mission host (this can be any team member, although in practice the team lead would normally take this role) and each other VBS client joins the mis- sion remotely. Our work to date requires manual selection of team roles by each member, but this could be automated in future iterations to automatically assign preset member roles. The VBS DIS (Distributed Interactive Simulation) data bus enables communication between all the client machines and allows the VBS clients to exchange information. In our architecture, this is the only communication occurring between client machines. That is, the GIFT instances themselves do not directly communicate to each other and do not exchange any data or measurements. Instead, they each independently compute measures and feedback based on the information captured from the DIS bus. This teamwork architecture requires development of only a single master DKF that computes all measures at the team level. Each team member’s GIFT client runs an instance of GIFT independently with this master DKF. This vastly simplifies the calculation and feedback of teamwork measures.

This approach of a single cloned DKF for all team members has several advantages. The main advantage comes in ease of training content development and deployment. The effort in developing DKFs can be considerable. Use of a single file can reduce the time required for a training content developer to create training packages and allow them to focus on measurement rather than file management. Cloning the GIFT scenarios across all team member machines is fast since there are no individual differences between them. In addition, the single DKF framework ensures all team members receive the same feedback at the same time (ignoring network lag) since all calculations are performed on the information flowing through the DIS bus.

However, there are limitations to this approach. Individualized measurements and feedback are not possible with this current architecture as they would require customization of individual DKFs. We believe, how- ever, that ongoing efforts in improving DKF development and management tools will enable easier cus- tomization of DKFs to individuals in the future. A hybrid approach may be possible where a single team DKF is used in conjunction with individual DKFs to provide both team-level and individual measurement and feedback. Since the effort described here is focused on teamwork measures, the current architecture provides sufficient flexibility at that level.

Team Training Scenario

The scenario developed for this effort consists of a number of distinct, but narratively related, events or vignettes that the team plays through. Although the map follows a mostly linear path, the vignettes are broken apart as independent training elements. This provides the ability to dynamically reconfigure the training to adapt to high or low-performing teams, helps facilitate repeated training so the team is not able to predict how the scenario unfolds, and enables the delivery of training feedback in GIFT at the end of each vignette. In addition, the vignettes each provide a number of opportunities to train a variety of team- work skills. In this section, we describe each of the vignettes, as well as the expected actions and protocols for the team. In the subsequent section, the teamwork measures for these vignettes are discussed.

The scenario is adapted from an Army Basic Leader Course (BLC) Combat Search and Rescue (CSAR) training scenario. The team consists of a nine-member squad: two four-person fire-teams and a squad leader. The scenario takes place along a linear path through a wooded area. In the scenario, an F-16 pilot has ejected and landed in the area and their medical condition is unknown. The primary objective is to perform a search and rescue for the downed pilot, with a secondary objective of reaching a small village after the pilot is rescued. The team receives intelligence that enemy militia are known to be in the area and are hostile to our presence. [Figure 2](#_bookmark1) depicts the scenario map and layout of each of the vignettes.

Vignette 1: Encountering an Improvised Explosive Device (IED)

Within the scenario, there are two vignettes (vignettes one and four) in which the team encounters a poten- tial IED threat. One is a hoax IED and one is real, although the team is not aware of the status of either. It is expected that they treat both as potentially deadly. The IEDs are represented as deceased dogs in the road, a known method of IED camouflage. Wires protruding from the animal provide further indication of the threat. Upon identification of the IEDs, the team member who first notices it is expected to halt and inform the rest of team of the threat. The team then coordinates their sectors of fire to ensure they have full 360-degree situational awareness of the environment. This requires coordination and communication within the team to secure and clear the area. Next, the team is expected to address the 5 “C”s of IEDs. These are Confirm, Clear, Call, Cordon, and Control. These choices are given through a menu system within VBS and the team is expected to choose the right “C”s in the right order.

Vignette 2: Finding and Rescuing the Pilot

As the team enters the area known to contain the downed pilot (the general area is provided in the pre- briefing material), they are expected to maintain formation and begin visually scanning the area. The pilot is located near the path, but hidden amongst the foliage. This requires the team to search the area. If not found in a set amount of time, the pilot will release a flare to draw the team to her location. Once the pilot is located, the team is expected to coordinate a number of activities including securing the pilot, setting up a cordon with interlocking sectors of fire to scan for threats, applying first aid, moving the pilot to an open landing zone, issuing a 9-line MEDEVAC request, and waiting for the helicopter to extract the pilot.



Figure 2. Scenario Map and Vignette Layout

Vignette 3: Reacting to Harassing Fire from Hostile Militia

At a point along the path, the team encounters harassing small-arms fire from two hostile militia members. This occurs suddenly and without warning, requiring the team to react quickly. In VBS the damage and accuracy of the hostile actors is reduced so as not to injure any of the team members (although this fact is not known to the team). The goal of this vignette is for the team to coordinate an immediate response to the threat without jeopardizing the other mission objectives. We require the team to contain the threat by returning fire to either kill or scare-off the hostile actors, but we do not expect them to chase after the militia. Specifically, the team is expected to take cover, return fire, call clear once the threat is neutralized, and continue on the path.

Vignette 4: Encountering an Improvised Explosive Device (IED)

The fourth vignette involves encountering a real IED (although the team does not know whether this one is real or another hoax as in vignette 1). The goals, actions, and measures for this vignette are the same as vignette 1, with the exception that if the team gets too close to the IED, it will detonate, ending the vignette.

Vignette 5: Reacting to Unknown Individual

In the final vignette, the team encounters an individual pulling a cart along the road. The individual’s identity and intentions are unknown. The individual is heading directly towards the team, and while not overtly threatening, he still poses a potential risk. The team is expected to quickly notice the individual, halt movement, communicate his presence and approach to the rest of the team, and “Show, Shout, Shoot,” that is, escalate fire at the individual until he retreats or is otherwise neutralized. The objective is for the team to react quickly, communicate the evolving situation to each other, and secure the individual (either through lethal force or until he leaves).

Team Training Measures

Measurement and assessment of teamwork skills within the training environment is crucial to ensure that the system is providing the correct opportunities for skill development, but also to provide the team feed- back on their teamwork training progress. The development of measures for this effort occurred largely in parallel with scenario development. Specific elements of the vignettes were chosen because they provide opportunities to assess teamwork skills. To develop teamwork measures, we use a process based on the Rational Approach to Developing Systems-based Measures (RADSM; Orvis et al., 2013), which has been successfully used to develop indicators and measures of team states (McCormack, Brown, Orvis, Perry, Myers, 2017). The RADSM process consists of several steps that ensure that developed measures are con- ceptually sound and contextually relevant. The end result of this process is a set of teamwork measures that can be assessed automatically and unobtrusively (that is, not requiring human coding or input) given the data available in the system. We describe this process in our previous GIFTSym paper (McCormack, et. al., 2018), so we focus here only on the resultant measures.

Selection of the specific teamwork construct to train was motivated a previous meta-analysis (Sottilare, et. al., 2017) in which a number of teamwork themes were identified, including coordination, cohesion, com- munication, cooperation, conflict, and others. The teamwork measures developed for this effort focus pri- marily on coordination. This construct is defined in various ways throughout literature, but for our purposes we treat coordination as the synchronization and awareness of team member actions in pursuit of a common team goal. Communication among team members is a large part of coordination (and most teamwork con- structs), but communications were not able to be captured and analyzed in GIFT during this effort. Verbal communication analysis requires the ability to capture individual speech utterances from each team member and perform speech-to-text processing on the audio. There is currently no approach integrated into GIFT to analyze verbal communications and it was beyond the scope of this effort to develop one. Furthermore, while the text chat (such as instant messaging apps or VBS’s built-in capabilities) offers an alternative solution, it was deemed untenable for this scenario. Typing messages would require participants to stop moving and performing actions within VBS, which would interfere with the often highly-kinetic vignettes. As such, our focus of measurement was not on the communications themselves, but on the actions and behaviors that require communication and coordination to occur.

We describe a selection of coordination measures developed for this scenario in [Table 1](#_bookmark2). Several of these measures repeat across vignettes (such as maintaining team formation) and others are omitted here for the sake of brevity.

Table 1. Example Measures of Coordination

|  |  |  |
| --- | --- | --- |
| **Measure Name** | **Description** | **Measure Feedback** |
| Maintain- ing Team Formation | During movement along the road the team should maintain a close formation. This initial measure will focus on the geodesic distance of the team, rather than on the relative positioning of each player. Geodesic distance is defined as the farthest distance any two team members are apart. That is, we measure the dis- tance between each pair of individuals, and take the maximum. The geodesic distance and the teams (above/at/below) expectation rating should be reported every 15 seconds. | Above Expectation: 8m < Geodesic distance < 12m |
| At Expectation: 5m < Geodesic distance < 8m OR 12m < Geodesic distance < 15m |
| Below Expectation: Geodesic distance < 5m OR Geodesic distance > 15m |
| Completing a Team Halt | The first person to notice a threat (IED, unknown indi- vidual in the road) should halt and give the command to teammates to halt. We define a maximum and min- imum distance from the threat. Once an individual is within the maximum distance radius, we monitor their movement and look for a stop of movement. We then monitor other team members to identify if/when they stop movement. If an individual has reached the mini- mum radius before a halt occurs, the entire team fails. | Above Expectation: first person halts between mini- mum/maximum distance, calls halt, other team members halt within 3 seconds of call. |
| At Expectation: first person halts between minimum/maxi- mum distance, calls halt, some team members halt within 3 seconds, but other team members halt before reaching mini- mum distance, but after 3 seconds. |
| Below Expectation: Any team member crosses minimum distance radius before halt is complete. |
| Completing the 5 C’s of an IED En- counter | After noticing an IED, the team correctly selects the “5 C’s” (Confirm, Clear, Call, Cordon, Control) from the drop down menu in the right order. Each person may only select one item from the list, and the selec- tion should be completed in a timely manner. Team members should communicate and coordinate on who is completing a selection and the correct order of se- lection. | Above Expectation: 5 different individuals correctly choose the 5 “C” s in the right order from the menu. Completes se- lection within one minute. |
| At Expectation: At least 4 different individuals choose the 5 “C”s, but in the wrong order or it takes longer than one mi- nute, but less than two minutes. |
| Below Expectation: The team fails to select the correct 5 “C”s, makes an incorrect selection, or takes longer than twominutes, or selections are only made by 3 or fewer individu- als. |
| Attaining Visual Control of the Envi- ronment | The team should scan their surroundings for potential threats. To measure this, we take each team members orientation angle and assume that they can visually as- sess a +-30-degree arc from their orientation angle.The union of all of these field of view (FOV) arcs across the team should cover 360 degrees. | Above Expectation: There are less than 10 degrees not cov- ered in the union of FOVs. |
| At Expectation: There are between 10 and 30 degrees not covered in the union of FOVs. |
| Below Expectation: There are greater than 30 degrees not covered in the union of FOVs. |

Table 1 (Continued). Example Measures of Coordination

|  |  |  |
| --- | --- | --- |
| **Measure Name** | **Description** | **Measure Feedback** |
| Identifying Hostile Actor Location | After the first hostile bullet is fired we measure the time it takes to identify the location of the hostile ac- tors. The timer ends when one of the criteria is ful- filled: at least one team member has spotted the hostile actors through the binoculars or at least half the team members have their weapons pointed within +-10-de- grees of the hostile actors. | Above Expectation: The team spots the hostile actors within 10 seconds. |
| At Expectation: The team spots the hostile actors within 20 seconds. |
| Below Expectation: The team takes longer than 20 seconds to spot the hostile actors. |
| Applying First Aid | After finding the pilot, the team correctly selects the MEDEVAC actions (Assess the pilot’s condition, Identify and Control Bleeding, Assess Breathing and Chest Injuries, Check for Burns, Monitor for Shock) from the drop down menu in the right order. Each person may only select one item from the list, and the selection should be completed in a timely manner.Team members should communicate and coordinate on who is completing a selection and which is the cor- rect order of selection. | Above Expectation: 5 different individuals correctly choose the First Aid steps in the right order from the menu. Com- pletes selection within one minute. |
| At Expectation: At least 4 different individuals choose the First Aid steps, but in the wrong order or it takes longer than one minute, but less than two minutes. |
| Below Expectation: The team fails to select the correct First Aid steps, makes an incorrect selection, or takes longer than two minutes, or selections are only made by 3 or fewer indi- viduals. |
| Taking Cover from Harassing Fire | After the first bullet is fired by the militants, we meas- ure the amount of time it takes for all team members to take cover. In game, we define taking cover as ei- ther lying prone or crouching. The outcome metric is the amount of time between the first hostile bullet fire and the last team member to take cover. | Above Expectation: The team takes cover within 5 seconds. |
| At Expectation: The team takes cover within 10 seconds. |
| Below Expectation: The team takes longer than 10 seconds to take cover. |

While the measures discussed here were initially developed with VBS in mind, we note that there is high potential for transfer to other learning environments. By abstracting away the specific doctrinal details, we arrive at general teamwork measures and GIFT condition classes that can be instantiated elsewhere. As an example, consider the “5 C’s” measure from the IED vignette and the first aid measure from the pilot rescue vignette. We refer to these as “team sequence” tasks, where team members must select the right choices, in the right order, across different team members. The developed condition class for this is applicable to any training environment where coordinated selection of ordered choices is required. Another example is the team formation measure, where team members are required to stay relatively close but not bunch up. The corresponding condition class measures the geodesic distance of the entire team and compares that against acceptable thresholds. This condition class is applicable to any spatially-oriented training environ- ment, where relative location of team members is important.

CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

The new architecture for training teams in GIFT discussed in this paper provides a straightforward, scalable approach to delivering teamwork skill training content. By utilizing one team-level DKF, content develop- ers can focus on measure development, implementation, and training delivery. While currently only team- level feedback can be delivered through this approach, individualized or sub-team feedback will be possible with future implementations.

The scenario developed under this architecture provides realistic team training in a virtual environment. The individual vignettes are reconfigurable and provide opportunities to develop a variety of teamwork skills. The initial focus of team training has been on coordination, but future work will expand to measure cohesion, communications, conflict management, and others. The new condition classes implemented in GIFT are generalizable to a variety of training scenarios and training content.

Future technology development will also seek to incorporate naturalistic communication assessment through speech-to-text and natural language processing capabilities. This will further enable teams to gain invaluable skills, while minimizing the need for more obtrusive or burdensome communication assessment techniques, such as chat or menu selections, which can distract from the central goals of the training. Fi- nally, additional customizations to GIFT will provide better immediate and after action review feedback through the use of the learner action panel.

ACKNOWLEDGEMENTS

The research reported in this document was performed in connection with contract number W911NF-17- C-0061 with the U.S. Army Contracting Command - Aberdeen Proving Ground (ACC-APG). The views and conclusions contained in this document/presentation are those of the authors and should not be inter- preted as presenting the official policies or position, either expressed or implied, of ACC-APG, U.S. Army Research Laboratory or the U.S. Government unless so designated by other authorized documents. Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwith- standing any copyright notation hereon.

REFERENCES

Army Doctrine Reference Publication (ADRP) 6-0, Mission Command. Washington, DC: Headquarters, Department of the Army, 2012.

Bonner, D., Gilbert, S., Winer, E., Dorneich, M., MacAllister, A., Kohl, A., et al. (2017). Military Team Training Utilizing GIFT. Paper presented at the Interservice/Industry Training, Simulation, and Education Confer- ence (I/ITSEC), Orlando, FL.

McCormack, R.K., Brown, T.A., Orvis, K.L., Perry, S., & Myers, C. (2017). Measuring Team Performance and Co- ordination in a Mixed Human-Synthetic Team Training Environment. Paper presented at the Interserv- ice/Industry Training, Simulation, and Education Conference (I/ITSEC), Orlando, FL.

McCormack, R., Kilcullen, T., Sinatra, A., Brown, T., & Beaubien, J. (2018). Scenarios for Training Teamwork Skills in Virtual Environments with GIFT. Paper presented at the Generalized Intelligent Framework for Tutoring (GIFT) Users Symposium (GIFTSym6), Orlando, FL.

Orvis, K. L., Duchon, A., & DeCostanza, A. (2013, January). Developing Systems-based Performance Measures: A Rational Approach. In The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC) (Vol. 2013, No. 1). National Training Systems Association

Salas, E., Cooke, N.J., & Rosen, M.A. (2008). On teams, teamwork, and team performance: Discoveries and devel- opments. *Human Factors, 50*, 540-547.

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

Sottilare, R. A., Burke, C. S., Salas, E., Sinatra, A. M., Johnston, J. H., & Gilbert, S. B. (2017). Designing adaptive instruction for teams: A meta-analysis. *International Journal of Artificial Intelligence in Education*, 1-40.

Wilson, K. A., Salas, E., Priest, H. A., & Andrews, D. (2007). Errors in the heat of battle: Taking a closer look at shared cognition breakdowns through teamwork. Human Factors, 49, 243–256

ABOUT THE AUTHORS

**Dr. Robert K. McCormack** is a Principal Mathematician and Deputy Director of the Performance Assessment Tech- nologies Division at Aptima. He has expertise in the areas of unobtrusive measurement, computational linguistics (NLP), machine learning, epidemiological modeling, and human sociocultural modeling and analysis. Dr. McCor- mack received a Ph.D. and M.S. in Mathematics from Texas Tech University, and a B.A. in Mathematics and Computer Science from Austin College

**Ms. Tara Kilcullen** is a Program and Customer Engagement Lead at Aptima, Inc. In her role, she supports business strategy, planning, and execution, defines market needs for technology requirements, product maturation and suc- cessful transition of science and technology (S&T) research programs, and leads defense programs that differentiate Aptima as an industry leader in Modeling, Simulation and Training. Ms. Kilcullen holds B.A.’s from the University of Pittsburgh and an A.S. from Full Sail University as well as several certifications.

**Dr. Anne M. Sinatra** is a Research Psychologist, and part of the adaptive training research team within CCDC Soldier Center – STTC’s Learning in Intelligent Tutoring Environments (LITE) Lab. She works on the Generalized Intelligent Framework for Tutoring (GIFT) project. Her background is in Human Factors and Cognitive Psychology.

**Mr. Alexander Case** is a Software Engineer in the Product Engineering Division at Aptima, Inc. His technical inter- ests focus on front- and back-end applications, data collection and analysis, and video game and simulation develop- ment. Other areas of interest include application telemetry, containerization, continuous integration, and configura- tion management. While at Aptima, he has worked on Angular web applications and developed server-side applica- tions in Python and C#. Mr. Case holds a Bachelor’s degree in Computer Science from Bridgewater State University.

**Mr. Daniel Howard** is a Principal Software Engineer and Product Line Lead for Data Management and Visualiza- tions at Aptima, Inc. Mr. Howard works on developing a variety of applications focusing on complex visualization and the analysis of large volumes of data, and is responsible for the efficient storage and retrieval of this measurement data through Aptima’s ASA™ framework. Mr. Howard holds a M.S. and B.S. in Computer Engineering from Roches- ter Institute of Technology.