1

User Feedback on a Hybrid Team Tutoring Strategy

Benjamin Goldberg [1], Jeanine A. DeFalco[1,] Michael Hoffman[2] and Elyse Burmester[2]

[1] US Army DEVCOM Soldier Center, Orlando, FL 32826, USA
[2] Dignitas Technologies, Orlando, FL 32817, USA

Abstract
In this paper we present the outcomes of a user centered qualitative usability evaluation across a set of tools and methods used to support a hybrid team intelligent tutoring strategy. Feedback was received across user tools designed to interface human trainers with adaptive instructional components used to monitor performance in real-time, promote reflection and discussion during a scenario review, and being able to explore performance and data over time to track competency development objectives. The methodology and results of the interviews are shared, with a discussion focused on insights and required features these tools require.

Keywords
Hybrid Intelligent Tutoring, GIFT, Team Tutoring, Competency

1. Introduction

In this paper we present work related to on-going research focused on the implementation of tools and methods to support team development in military domains. Specifically, we will present and discuss the outcomes from a recent set of structured interviews with defined stakeholders centered on the utility of adaptive instructional technology supporting their training needs. The interviews were administered to collect user-centered feedback and to better define interaction requirements that need to be accounted for in development and implementation.

This research is directly informing the Synthetic Training Environment Experiential Learning for Readiness (STEEL-R) [1] project that is producing an extensible data strategy for mapping performance outcomes derived from a simulated environment across persistently tracked competency frameworks. This is being accomplished by integrating the Generalized Intelligent Framework for Tutoring (GIFT) [2] with functional components and data standards in the Advanced Distributed Learning (ADL) Initiative's Total Learning Architecture [3].

Through this approach, we are establishing an architecture that supports a distributed layering of data inferencing. This chain of processing converts raw data into metrics during run-time, applies those metrics to assess performance under varying contexts, and reports those assessments to update long-term competency representations. This approach is being applied within a team training use case, with a focus on modeling and impacting the development of competencies related to task roles and effective team behaviors.

2. A Hybrid Team Tutoring Strategy

In the context of STEEL-R, it is recognized that the intelligent tutoring functionality is intended to optimize observers and trainers managing a learning event, rather than to replace them. The goal is to

use intelligent tutoring services where appropriate to offset the workload and task load performed by instructors, and using data driven methods to build objective assessments where possible. Part of this effort is to establish user-friendly devices that interface adaptive instructional services with humans across all phases of a learning event [4]. For this workshop, we are sharing the outcomes of structured interviews held with individuals of an operations group for an active Army unit. We presented the current state of the GIFT Game Master Interface, and mock-ups of a new proficiency and competency dashboard. For this context, we are examining three levels of engagement, which include a run-time execution tool, an after-action review (AAR) playback tool, and a persistent competency dashboard tool. A high-level description of each tool will be provided, followed by a breakdown of the interview outcomes.

2.1.   Training Execution Tool for Observer Controller/Trainers
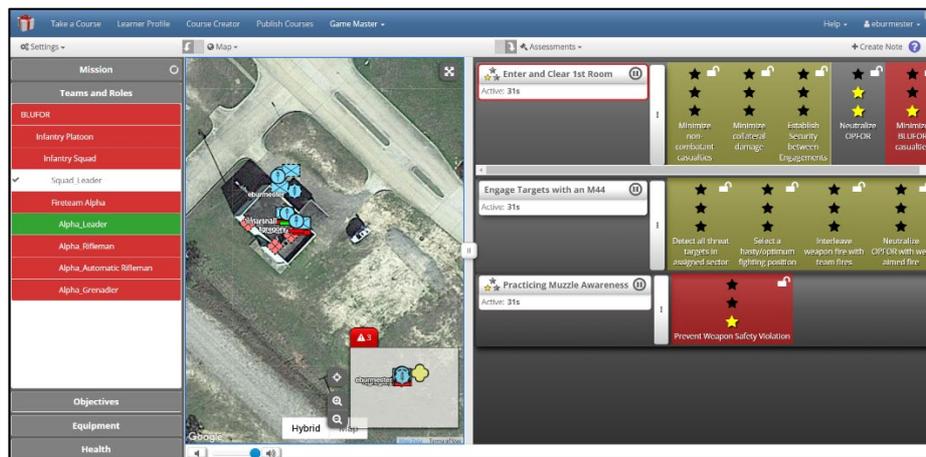
Figure 1: The STEEL-R Game Master interface used during run-time with visualization of the interaction space and underlying assessment structures.

This set of User Interfaces (UIs) are used to support the trainer during the real-time delivery of training. The tool connects with available data sources and the GIFT domain module, which provides an overview of real-time performance as tasks and scenarios are completed. The tool also provides an ability for the human observer to insert assessments and create bookmarks for display during AAR.
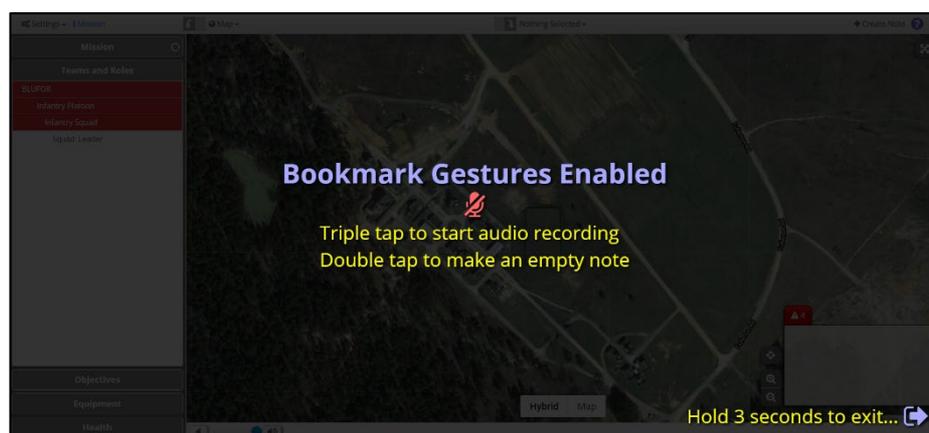

Figure 2: The Game Master run-time mode with the Bookmark Gesture function enabled.

2.2.   Adaptive AAR and Scenario Playback Tool
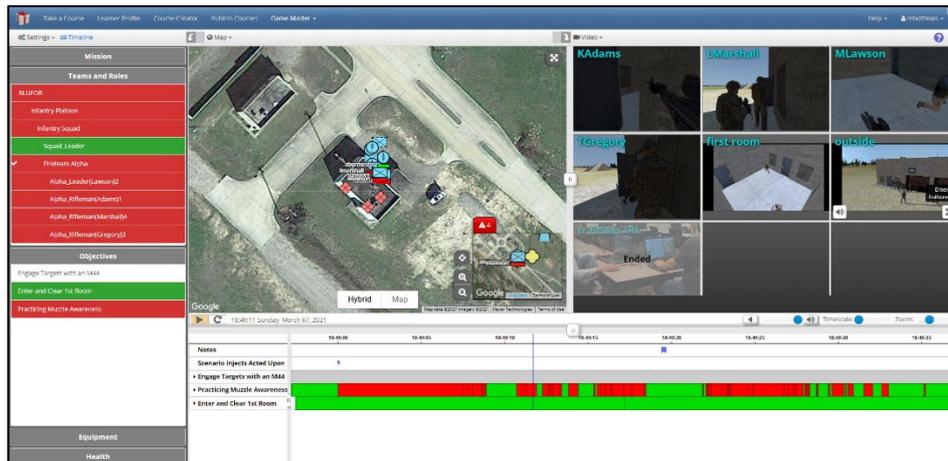
Figure 3: The Game Master in playback mode with a previous log visualized for exploration and AAR.

This set of UIs are used to support an instructor or team leader with reviewing performance outcomes following completion of a scenario. It provides visualizations and dashboards based on aggregated data to help drive effective AARs. The tool displays all the measures of performance and effectiveness collected during training on a searchable timeline with play-back capabilities. A user can explore specific tasks, re-play scenario events, and highlight observed assessment bookmarks that require attention and reflection during the AAR. Coaching strategy recommendations will be added in the near-term.

2.3.   Dashboard for Competency and Readiness Tracking

This set of UIs and dashboards are being designed with team leads and commanding officers in mind. The tool will visualize experiential data collected across several instances of training. This can involve examining performance across multiple scenarios performed in a single day (i.e., to examine performance trends as they relate to that day's specific training objectives), or across multiple training events and environments experienced during a training cycle (i.e., to track competency development over time and to monitor readiness requirements). These visualizations are based on stored xAPI (eXperience Application Interface [2]) statements that are collected over time and across multiple environments. Through these interfaces, a user can explore how their team and sub-ordinates are progressing in the development of operational competencies.
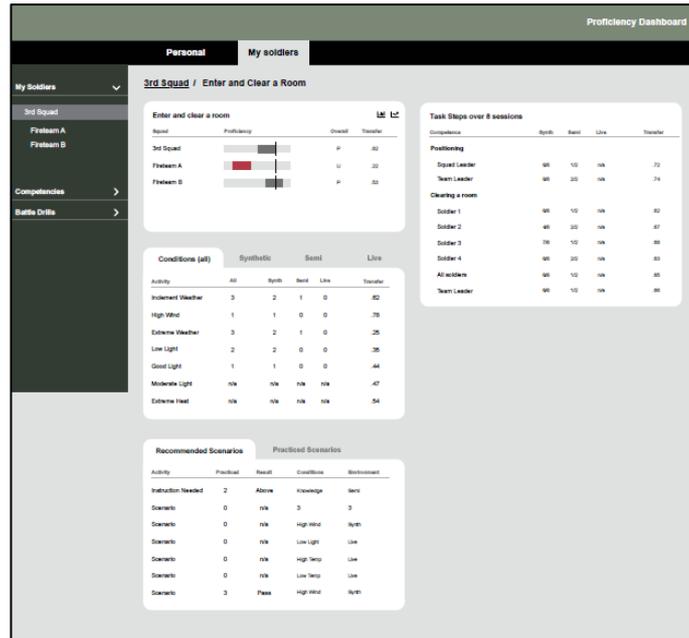
Figure 4: STEEL-R draft readiness dashboard that visualizes long-term competency tracking information based on persistent data modeling.

3. Methodology

To date, the UIs described above have been developed with input from limited retired Subject Matter Experts. In an effort to garner explicit feedback from the real end-user community on layout, features, data types, and data formats, a series of interviews with designated Squadron-level personnel were executed. These interviews collected think-aloud feedback while each tool was demonstrated from a functionality standpoint.

3.1. Participants

The tools were demonstrated to and reviewed by 3 groups of participants averaging 2-6 participants per group. Participants included members of the Operations Team for an active Army unit, whom are responsible for coordinating and supporting training planning and delivery at the squadron level. This group consisted of former squad and platoon leaders, troop commanders and squadron leadership. These individuals also serve as Observer/Controller Trainers (OC/T) who perform oversight and assessment services while training is delivered.

3.2. Procedure

Each interview followed the same sequence of interactions. A short briefing was provided to introduce the technology and the STEEL-R program the research is supporting. This was followed by an active demonstration across the three hybrid tutoring interfaces described above. The demonstration showed the recording of a simulated squad executing Battle Drill 6a, Enter and Clear a Room, in a synthetic game environment. During this portion, exercise control capabilities available to the OC/T were highlighted, as well as a review of the objective-associated assessments established using GIFT's run-time capability.

This was followed by a demonstration of the AAR playback experience for that exercise. The stored log-file was loaded and the participant group was able to explore the visual playback functions that are utilized for AAR and reflective discussions. After the AAR demonstration, participants then reviewed

the readiness dashboard that was designed to visualize competency development overtime based on persistent data capture.

Two scenarios were presented to the participants: Scenario 1 demonstrated novice performance, and Scenario 2 presented journeyman level performance for the same task. For both scenarios, participants were walked through the following interactions and workflows:

- Training Execution Tool
  a. Using the Game Master tool to assess squad training IN-ACTION.
  b. Input an assessment for a task that had not been assessed.
  c. Filter assessments for an individual or sub-team.
  d. Insert a text bookmark using the gesture input.
  e. Identify squad performance for specific Knowledge, Skill, Ability components.
- Adaptive AAR and Playback Tool
  a. Switch from active session to playback mode for AAR.
  b. Use Game Master tool to review/edit recorded performance and prepare AAR.
  c. Edit text bookmark created during scenario execution.
  d. Use Game Master to execute an AAR.
  e. Identify a bookmark created during scenario execution.
  f. Identify scenario injects applied during scenario execution.
  g. Describe squad performance across "Muzzle Flagging" task.
- Competency and Readiness Dashboard
  a. Log-in to track team development towards readiness.
  b. Examine competency trends and performance levels across all critical tasks and skill components.
  c. Determine training needs at the individual and team level based on visual breakdown of competency levels across role and team dimensions.

During the demonstration, participants answered face-to-face loosely structured interview questions that elicited free response and discussion. Participants were also able to interact with the Game Master active and playback modes. These discussions were video recorded and transcribed at a later date.

3.3. Data

The data collected from this engagement is primarily qualitative feedback based on the interviews. Participants responded to a series of questions focused on these themes:

- Evaluation and critique across each interface component.
- Compare current training review processes to hybrid team tutoring workflows.
- Assess likelihood and challenges of using tools during training.
- Describe current training event data vs proposed methods.
- Recommend areas for improvement focus.
- Provide overall feedback on hybrid team tutoring concept.

The data was used to establish requirements and notes for iterative improvements to the interfaces in question. It also highlights prioritization of functionality based on the goals and needs of the training audience.

4. Results

Video logs across each interview group were transcribed and a thematic analysis was conducted. Overall, there were three primary themes addressed across all workflows:

1. Positive Feedback

A short synopsis of the key takeaways will be provided, followed by a conclusion section on planned changes to be made as a result of this activity.

4.1. Positive feedback

Participants noted that the bookmark function was "great" and the capability to revisit AAR data at a later date was valuable. The gesture bookmark function was noted as a critical feature, as it would allow uninterrupted performance annotations of ongoing training. **Participants** also appreciated being able to see individual Soldiers on one side of the screen and the big picture of the training event on the other side of a screen. This was noted for both active and playback mode within the Game Master

Soldier 2 -- "This system is awesome; the technology is really cool; I think it's a great asset."

Compared to how AARs are currently managed, the Game Master tool was assessed as definitely having value.

Soldier 3 – "For AAR purposes, this would be a great tool to speed up the process. With this tool, we predict Soldiers would pay attention in AAR and would benefit from actually seeing their interactions in training through a playback function."

Soldier 4 – "As far as the AAR stuff goes talking about it is it's a bright system it's pretty straight forward, it's not overwhelming: video, tracker, roles – it's blunt and to the point. I do like that about the system. Based on how AARs are done right now, this would definitely be valuable."

Participants also thought that the Game Master tool would speed up the AAR process, and that trainees would benefit from seeing their actions in training through the playback function. Overall, there was positive feedback regarding the AAR feature that it would not only capture errors and bad habits, but also allow **participants** to provide positive feedback to trainees, reinforcing the good choices and not just focusing on the negative.
There was also good feedback related to an associated competency dashboard. All groups found it critical to track progress over time, but currently there are no robust methods to do so. While the concept is positively viewed, they highlight human factors dependencies that must be fully accounted for to support adoption of this tech.

Soldier 3 – "I think data matters from a macro perspective, but for the team lead, they're going to know that their individual team members have these issues and what they're going to work on."

4.2. Critique of System

Participants emphasized the importance of simplicity and ease of navigability across all tools. Participants thought assessment explanations should be easier to access and their meaning more obvious. In terms of the Assessment Panel, feedback included that the panel should show concepts with automated assessments when something goes wrong; that assessment explanations should be easier to access and more obvious; and that the assessment panel should default to show only observed assessments, with an option to show full display.

Soldier 4 – "My only concern is if it's focused on a single echelon, there needs to be a way to quickly swap to other team representations."

Soldier 1 – "Unlikely that the observer controller will be engaged with this tool during run-time."

Soldier 5 – re: proficiency ratings on competency dashboard "it would be easier if it said like needs more training or specific 3 scenarios, something other than just a color bar."

While the user interface tools and methods were mostly well received, these associated critiques will be carefully examined for critical requirements to associate in the next iterative improvement cycle. The primary critique centered on human factors and usability, highlighting how important the interfacing components will be to the success of these training management methods by those who will use them. Additional information collected from the interviews was also examined from a tool requirement and dependency viewpoint.

4.3.   Tool Requirements and Dependencies.

Participants noted that their primary concern in training management is enforcing safety and ensuring a value-added outcome is experienced:

Soldier 1 -- "Safety is primary concern, flagging, weapon safety, overall risks of terrain, then overall training value, is the training effective, are they getting anything from it, having the Team Leads engaged, reporting what their team was supposed to do"

Soldier 4 – "Spec out whatever the leadership says are important."

These are the elements an OC/T is focused on during a learning activity, and any supporting tool must account for these requirements. Currently, the majority of performance is collected via a clipboard and an itemized list of performance steps that receive a binary 'go' or 'no-go' grading. Based on the goals of maintaining safety and eyes on the training, these performance annotations are often input after a training scenario is complete, and is dependent on an OC/T accurately interpreting and recalling their observations. Interfacing adaptive instructional system technology should be introduced from a user-goal standpoint, and establishing workflows and support features that produce a viable training experience with retained knowledge and skill.

Regardless of the tools and techniques applied, the interface must be intuitive and supportive of training workflows, and consistent across all modes of learning the technology is implemented within. It must also provide intuitive visualizations of data to ensure accurate and actionable insights are derived. A big focus here is interfacing common adaptive instructional system services with human counterparts and exposing relationships and logic to help drive skill development.

Soldier 2 – "For real-time, during actual training, good likelihood of clicking the wrong thing, so interface has to be super basic."

Soldier 3 – Re: how performance measures are visualized on game master: "Request for more explicit color coding and meaning of stars."

Soldier 5 – re: adding adaptive logic to the game master: "Function to add stressors based on observed performance."

When comparing the three tools listed above, during execution of a scenario, the tool must be a secondary focal point at times of task completion. It will be critical to promote the use of bookmarks and audio annotations to capture in-situ observations and assessments that can be accounted for at a later day. The primary interest is in using this technology during an AAR and for long-range performance tracking. However, it will be important to identify a strategy for getting users of the tool to build labeled data sets of observed performance for the purpose of information machine learning classifiers in support of automating the assessment space.

5. Recommended Changes based on Feedback

A significant takeaway from these interactions is a necessity to focus on simplicity. There is often an emphasis to include all features of a tool to a user so that every component and service is made available. For the successful utility of a hybrid-tutoring approach, understanding what features are necessary and what features are nice to have at each step in the workflow will dictate a forward focused design. We will adjust the current default model to visualize the basic assessment and bookmarking functions to the OC/T, while providing quick shortcuts to all other associated function.

Quick-win recommendations from participants are captured in Table 1. These include having the bookmark audio button always present on the task/concept panel, use visual and auditory notification when a task ends, and better correlate Domain Knowledge File (DKF) places of interest on a rendered Google map. In addition, participants recommend having a widget that would allow for an easy switch between pre-defined perspectives with an option to customize each; a simple touch screen that in addition to the global book gesture would include the ability to show or hide the assessment panel in a new User Experience (UX); and an advanced touch screen that would include elements of a simple touch screen but would show or hide assessment or scenario injects in a new UX.

Table 1.
UI Modifications and Recommendations

Tool		Action
Game Master	Assessment Panel	Always show bookmark record audio button on edit task/concept panel.
	Assessment Panel	Notify user when a task ends, and it contains one or more observed assessments without an assessment.
	Map	Show DKF places of interest on Google maps (points, lines, areas).
	General	Widget to easily switch between pre-defined perspectives (with option to customize each).
	General	'Simple Touch Screen': Current global bookmark gesture + ability to show/hide assessment panel in a new UX.
	General	'Advanced Touch Screen': 'Simple Touch Screen' + ability to show/hide assessment/scenario injects in a new UX.
Competency Dashboard		List date of last training for each task/competency
		Establish ranking scheme for comparison and motivation
		Display bio-data/stress overtime separate from performance
		Create easy report export for leadership reporting

While the technology seems welcomed by those we engaged with, there was still skepticism on the utility of these tools being widely adopted. Integrating new technology with new workflows/user roles requires careful coordination with viable stakeholders. This needs to be an iterative process, but focused more than just on Soldier feedback. With a "matured" version of each tool with touchpoint input incorporated, future work will require controlled studies to evaluate true impact and use characteristics.

6. Acknowledgments

7. References

[1] Goldberg, Benjamin, K. Owens, K. Gupton, K. Hellman, Follow the Data: From Raw to Credential". Presented at 2020 iFest. (2020).

[2] Walcutt, J.J., S. Schatz (eds.), Modernizing Learning: Building the Future Learning Ecosystem. Washington, DC, Government Publishing Office (2019).

[3] Sottilare, Robert, B. Goldberg, K. Brawner, H. Holden, Modular Framework to Support the Authoring and Assessment of Adaptive Computer-Based Tutoring Systems (CBTS). In Proceedings of the Interservice/Industry Training Systems & Educ Conf, Orlando, FL (2012).

[4] Goldberg, Benjamin, M. Hoffman, A. Graesser, Adding a Human to the Adaptive Instructional System Loop: Integrating GIFT and Battle Space Visualization. In: Design Recommendations for ITS, Vol. 7: Data Visualization, U.S. Army DEVCOM Soldier Center.