An Evaluation of the Generalized Intelligent Framework for Tutoring (GIFT) from a Learner’s Perspective

by Robert A Sottilare, Trent J Reece, Michael W Fung, James F McAuliffe, and Michael D Matthews
NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position 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.

Destroy this report when it is no longer needed. Do not return it to the originator.
An Evaluation of the Generalized Intelligent Framework for Tutoring (GIFT) from a Learner’s Perspective

Robert A Sottilare
Human Research and Engineering Directorate, ARL

Trent J Reece, Michael W Fung, James F McAuliffe, and Michael D Matthews
United States Military Academy
An Evaluation of the Generalized Intelligent Framework for Tutoring (GIFT) from a Learner’s Perspective

Cadets participating in this evaluation were assigned tasks related to the learner’s perspective as part of a course assignment in PL488E, an engineering colloquium. Their thoughts are shared herein along with technical challenges identified by the US Army Research Laboratory based on cadet observations.

Current US Army standards for training and education are group instruction and classroom training also known as one-to-many instruction. Recently the Army has placed significant emphasis on self-regulated learning methods to augment institutional training. Per the Army Learning Model, Soldiers will be largely responsible for their own learning. One-to-one human tutoring has been shown to be significantly more effective than one-to-many instruction, but it is not practical to assign each Soldier a personal mentor. An alternative to one-to-one human tutoring is one-to-one computer-mediated tutoring using artificially Intelligent Tutoring Systems (ITSS), which have been shown to be effective in promoting individual learning in static, simple, well-defined domains (e.g., mathematics). To be practical, high authoring costs and limited adaptiveness barriers must be addressed. This document describes the outcomes of an evaluation conducted at the US Military Academy to determine initial usability of the Generalized Intelligent Framework for Tutoring, a tutoring architecture constructed with the goal to reduce time and skill needed to construct ITSs while increasing their adaptiveness or ability to act autonomously to optimize user learning.

Self-regulated learning methods, intelligent tutoring system, Army Learning Model, adaptive tutoring, authoring tools, usability

Security Classification of: Unclassified Unclassified Unclassified

Limitation of Abstract UU

Number of Pages 20

TelephoneNumber (407) 208-3007

Robert A Sottilare

Report Approved for public release; distribution is unlimited.

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18
# Contents

## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>iv</td>
<td></td>
</tr>
</tbody>
</table>

## 1. Introduction

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

## 2. Evaluation of GIFT from a Learner’s Perspective

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Introduction to Learning in GIFT</td>
<td>2</td>
</tr>
<tr>
<td>2.2 Evaluation of GIFT User Interface</td>
<td>3</td>
</tr>
<tr>
<td>2.3 COIN Operations</td>
<td>5</td>
</tr>
<tr>
<td>2.4 Logic Puzzle Tutorial</td>
<td>6</td>
</tr>
<tr>
<td>2.5 Tactical Combat Casualty Care Game-Based Tutorial</td>
<td>8</td>
</tr>
</tbody>
</table>

## 3. Future Directions

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Human Interaction</td>
<td>10</td>
</tr>
<tr>
<td>3.2 Application of Cognitive Theories in Tutoring</td>
<td>10</td>
</tr>
<tr>
<td>3.3 Narrative</td>
<td>10</td>
</tr>
</tbody>
</table>

## 4. References

<table>
<thead>
<tr>
<th>References</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td></td>
</tr>
</tbody>
</table>

## Distribution List

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>
## List of Figures

<table>
<thead>
<tr>
<th>Fig.</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>List of folders and files within the first file</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>Scripts folder</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>COIN explained within GIFT</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>Pedagogical stage: user chats between learner and AI</td>
<td>6</td>
</tr>
</tbody>
</table>
1. Introduction

The standard for acquiring knowledge in institutional training within the US Army is split between traditional classroom training and live training. These methods are used to test recall and allow Soldiers to apply and test their skills respectively in varying conditions and against a set of standards. Over the past 30–40 years, virtual simulation has been added to the training toolbox and a debate has raged about what mix of live and virtual training is optimal. To augment institutional training and provide flexibility and accessibility for Soldiers who need training, the Army has recently emphasized self-regulated learning where Soldiers are largely responsible for managing their own learning. From a common sense point of view, it might not seem practical for each Soldier to be able to manage their learning without some guidance. This guidance, also referred to as coaching, mentoring, or tutoring, is usually provided one to one by a human tutor. Generally this function has fallen upon noncommissioned officers. However, the success of one-to-one tutoring recognized by Bloom (1984, 2σ effect size) and VanLehn (2011, 0.8σ effect size) is impractical to implement in large organizations like the Army.

Once we decide to pull the human tutor out of the instructional loop, our alternative is to provide one-to-one computer-guided instruction using Intelligent Tutoring Systems (ITSs), which have been shown to be effective in promoting individual learning in static (e.g., desktop), simple, well-defined (procedural) domains (e.g., mathematics, physics). Well-defined domains generally have one solution to a problem presented whereas ill-defined domains may have multiple paths to success. ITSs are a practical alternative to one-to-one tutoring but are costly to author (develop) and do not have sufficient adaptability to support more dynamic, complex, ill-defined domains represented in many Army operations. To address the needs of learners, authors, and analysts/researchers who use or might use adaptive tutoring technologies to learn, develop new ITSs, and analyze the effect of ITS technologies, the US Army Research Laboratory (ARL) created the Generalized Intelligent Framework for Tutoring (GIFT, Sottilare et al. 2012).

GIFT is a prototype open-source, service-oriented, adaptive tutoring architecture targeted to support automated authoring, automated one-to-one and one-to-many guided instructional experiences, and evaluation of effect to determine the impact of current and emerging tutoring technologies with regards to learning outcomes. Ultimately GIFT will be a community development project. Currently there are about 400 users in 30 countries who are registered users of GIFT, which is freely available at www.GIFTtutoring.org.

This report is 1 of 3 evaluating the usability of GIFT from 3 perspectives: learners, authors, and researchers/analysts. This report is focused on the learner’s perspective, which is about what people who use GIFT to learn new knowledge and skills think about their experience and the ease of use of GIFT in facilitating and managing their instruction.
The instructional management construct within GIFT is targeted at providing automatic instructional experiences that are tailored based on the changing needs of the learner and providing broad accessibility for mobile learners. Instructional management is largely in 4 themes: 1) meta-cognition and self-regulated learning, 2) affect, engagement, and grit, 3) guided instruction and scaffolding, and 4) natural language and discourse. The default instructional engine for GIFT is the engine for management of adaptive pedagogy (eMAP), which guides instruction through the selection of feedback and content based on the learner’s performance or understanding of the concepts under training.

This report outlines learner evaluations conducted by cadets within the Engineering Psychology Program, a part of the Behavioral Science and Leadership Department at the US Military Academy (USMA) as part of their coursework in “Human Factors of Military Training Simulations” (PL488E) during the Spring semester of 2014.

2. Evaluation of GIFT from a Learner’s Perspective

This section addresses usability of GIFT from an individual learner’s perspective. GIFT as it is designed today does not fully address instruction of teams or collaborative learning. Research is ongoing at ARL to begin the modeling and assessment of teams across a variety of training tasks and domains. The goal of adaptive tutoring research (under which GIFT was created) is to discover, innovate, and transition effective tutoring tools and methods to support adaptive, self-regulated instruction.

2.1 Introduction to Learning in GIFT

For the learner evaluation, GIFT manages media content presentation to the learner. The information presented (media) is sequenced by the domain module, which also includes libraries the tutor uses to provide feedback. From there information is transferred to the learner through the GIFT tutor-user interface (TUI) and/or a separate browser window for simulations, slideshows (e.g., PowerPoint), web pages, or video content. During instruction, GIFT evaluates how well the learner understands the information and modifies the instructional methods as needed to tailor instruction for the student and optimize the learning experience (Sottilare et al. 2012).

The instruction may also include a check on learning with selected questions at any time during the training. Based on the results from these questions, the pedagogical module recommends which instructional strategy (e.g., feedback, request for reflection) should be selected next. Once the strategy is recommended, the domain module selects an appropriate tactic (action) based on the strategy recommendation and the context (e.g., where the learner has progressed in the course, lesson, or concept (Sottilare et al. 2012).
Research has shown that using a media presentation program can be an effective method of instruction (Mayer and Johnson 2010) that promotes engagement and learning. Specifically, a study done at the University of Texas demonstrated that the use of PowerPoint slides and an interactive video result in increased performance (Bartsch 2003; Zhang et al. 2006). From this experiment they found that students enjoyed PowerPoint slides significantly more than static media such as transparencies. Further, the results illustrated that the students actually learned more from PowerPoint than from the transparencies. Finally, the study showed that quiz performances were significantly higher with the PowerPoint slides than the transparencies.

In another study at the University of Arizona, results indicated that interactive videos enhance learning effectiveness (Zhang et al. 2006). This study compared students’ learning through an interactive video, regular video, and no video (e.g., traditional classroom lecture). Following one of the instructional periods, they took a test on the topic that included questions more in-depth than the pretest. The results from this experiment suggested that students learned significantly more from an interactive video than from all the other learning environments assessed. From these 2 experiments, media has been shown to produce better learning outcomes than other, more traditional methods.

Mayer (2005) demonstrated specifically how multimedia is an effective tool for learning. Based on cognitive theory, multimedia helps to break down information so it can be effectively processed by the brain (Mayer 2005). These specific steps within the brain include processing information in the dual channels of visual and auditory, each channel having a limited capacity, and a coordinated set of cognitive processes are carried out during learning. Mayer helps to explain why, specifically, multimedia (e.g., video and PowerPoint) is an effective tool. GIFT is designed to take advantage of these results by attempting to increase the performance and learning of the user through the use of interactive media in an automated software package instead of more-traditional teaching methods.

2.2 Evaluation of GIFT User Interface

One of the most important components, if not the most important component, of any computer software program is the interface. The interface connects the user to the information being presented (e.g., content and feedback) and can determine the level of enjoyment the program provides and whether users will continue or discontinue its use. Designing an easy-to-use and easy-to-understand interface should play an integral role in the overall design of a program. Without an effective interface design, the program will ultimately not achieve its desired effects with users.

While GIFT is in the early stages of development, its interface and start-up procedures require improvement before it can become the premier tutor-building program that it has the potential to be. After the user downloads and installs the GIFT software, he or she is presented with a folder filled with numerous folders and dozens of batch files. To the inexperienced user it can quickly become overwhelming and confusing when trying to determine such a simple task as to start the
program. Within the first folder, there is a “launchGIFT” batch file (Fig. 1) that is supposed to launch the program and turn on all of the various modules of the program, sending the user to the Internet domain-based welcome screen where he or she can then access the various courses/lessons and tools that the program offers.

While this batch file is meant to operate as the Start Up executable, it may not always successfully start all the modules, forcing the user to open the Task Manager, terminate the program, and restart. When this is the case, the user has to reopen the initial folder, open the GIFT folder, and then open the scripts folder, and then execute the launchActiveMQ batch file (Fig. 2). After this file is finished running, the user then has to open the launchMonitor batch file. This file will then open the program monitor, which will launch all of the necessary modules and allow the user to access the Tutor Interface.

As has been described, this process of simply turning on the program can be very difficult and frustrating to the inexperienced user. Based on this evaluation, a simpler launch procedure has been implemented and is much simpler. Future versions of GIFT will implement as a specific learner interface. Other unique interfaces being developed are targets for “authors” and “power users” or researchers). Access to and manipulation of information within GIFT is to be managed
by these user-unique interfaces. After the initial evaluation of GIFT start-up procedures and its interface, we continued by evaluating a few of the courses/lessons currently being offered for use with GIFT, including tutors for counterinsurgency (COIN) operations, logic puzzles, and game-based tactical combat casualty care (TC3) training.

### 2.3 COIN Operations

When the user first opens the COIN Operations tutorial, an external PowerPoint presentation appears on the screen with an interactive play button in the lower right hand corner that allows the user to navigate through the slides. This specific tutorial was published in 2009 so it is one of the earliest tutorials to be published using the GIFT platform. Its purpose is to guide the user through a lesson on COIN operations, first defining and describing it with PowerPoint and then asking the user a series of reflective questions as a check on knowledge using an interactive chat with artificial intelligence (AI).

As a whole, the presentation very effectively communicates what COIN operations entail because it visually depicts and describes the relationship between all actors, including the host nation, coalition forces, civilian population, insurgents, and external state actors (see Fig. 3). However, this simple method of using only a slide presentation is limited in effectiveness because it relies solely on text and images. Improvements might be realized through the addition of more interactive media and assessments of learners during problem solving. This highlights a limitation imposed by authors. GIFT has the capability to integrate and sequence all types of media but it is up to the author what is and is not included in a course or lesson.

![Counterinsurgency Explained](image)

Fig. 3 COIN explained within GIFT
After the learner has completed the PowerPoint presentation, GIFT closes it and a separate chat window opens up in the Internet browser. The AI within the chat prompts the user to answer a series of questions on the lesson that he or she just went through. This is the pedagogical stage of the lesson. The chat starts off with GIFT asking, “Based on what you have read, what situation does the United States face?”, to which the user is supposed to respond using some of the information that he or she just learned. The user then responds, and the AI attempts to analyze what the user has said, as shown in Fig. 4. If the user responds with a correct answer, the AI acknowledges it, but if the user’s answer is not fully developed, the AI prompts the user to continue to explain by giving the user some hints on to what the answer is. This module is extremely effective in helping the user not only understand the information but also be able to coherently explain some of the concepts that he or she was taught during the lesson.

Fig. 4  Pedagogical stage: user chats between learner and AI

2.4 Logic Puzzle Tutorial

The Logic Puzzle tutorial is a very straightforward program that seeks to teach users how to solve a logic-based puzzle. This is achieved in 3 stages. First a somewhat interactive tutorial is presented in PowerPoint that explains the purpose of the program, breaks down the process, and instructs users how to best solve a logic-based puzzle. The second stage consists of a series of questions that act as a check on learning and are asked within the GIFT TUI rather than PowerPoint. The last stage consists of another PowerPoint slideshow that asks users to apply their learned skills to solve a full-length logic puzzle on their own. Before and after each stage, a survey is administered to determine the mood and confidence of the user regarding their ability to solve a logic puzzle. The critical analysis of the usability is vital in determining whether the interactive tutor is being implemented in a proper fashion.
The first criteria for usability that will be examined are the timeliness and visibility of feedback throughout the program. Overall, feedback is immediate and responsive to the user’s decisions. For instance, answering a question incorrectly during the first stage engages the tutoring system that will explain why your answer is incorrect. Then the system gives you an opportunity to input the correct answer. This is useful in pointing out the user’s mistakes during the tutorial without incurring any sort of negative consequences. The second stage asks questions within the GIFT TUI, but there is no help with the questions, and the question box merely repeats verbatim the instructions for answering the questions. This effectively leaves the user with only the knowledge retained from the first stage to answer these questions. Removing feedback is done purposely to allow the user to dig deeply to recall the process needed to solve the puzzle, thereby imparting deeper learning.

There is no feedback as to whether or not the user is answering the questions correctly during this stage. Many of the questions allow the user to check as many boxes as apply but then there is no feedback to identify if the user actually answered the question correctly. The issue with the feedback within the second stage is that there is no mechanism to determine how far the user has progressed within the puzzle. The final stage suffers from similar feedback issues as the second stage. There is no obvious mechanism by which to get assistance in interpreting the clues or solving the puzzle. In addition, there is no feedback when users submit their solution as to whether the solution is correct or not.

To fix these issues with feedback would require a few changes to the GIFT programming. First, during the second stage the feedback given should refer back to the instructions given during the tutorial. This ensures that the user receives the same instruction that they received before but this time in a much more applied context. Connecting the tutorial’s learning points with the applied questionnaire will reinforce learning. A system that tells the user the total number of questions and how many questions they have left in the applied stage would be useful. Finally, some sort of system that tells the user during the second and third stages whether or not they answered a question correctly, and how to correctly answer the question in the future, would be very useful in training users how to properly answer logic puzzles.

The generation of feedback, linkages to previous knowledge, and presentation of user progress are issues being addressed in the development of eMAP, the pedagogical engine driving decisions within GIFT. eMAP is based on Merrill’s Component Display Theory (CDT) (Merrill et al. 1992), which states that the best method to guide learning is to present rules (tell), then demonstrate successful examples (show), then test the user’s knowledge (ask), and, finally, allow users to apply their knowledge during practice (do). Implementing CDT in eMAP will allow GIFT to effectively manage feedback and build new knowledge based on old knowledge and experience while tying performance errors by the user to previously learned knowledge. Future research is addressing the issue of an open learner model that will allow users to see their progress and make adjustments to optimize learning outcomes.
The second usability criterion measured is the aesthetic design of the final product. The first stage of the tutorial consists of incredibly word-heavy PowerPoint slides. The large volume of words is distracting, making it difficult to discern what information on the slide is important and what is not important. To alleviate this, the design of the slides makes use of PowerPoint bullets or numbers to organize the information and highlight in red font the information that the tutorial deems important. The second and third stages have a relatively simple design, with the important information either highlighted or put in a box to show the importance of that given material. Other than the first stage, the only other issue with the Logic Puzzle tutorial’s aesthetic design is that the surveys do not make the best use of the space given them. Because the surveys are given within the GIFT TUI, there is not much room provided, and the number of words used in the surveys force the user to scroll from side to side occasionally to properly answer the questions.

To improve the aesthetics of the logic puzzle tutorial, 2 changes need to be made. First, the first stage of the tutorial needs to be revamped to reduce the amount of words placed on the PowerPoint slides. Better alternatives would be the use of symbols to represent words or having the user interact with the logic puzzle and then have an explanation pop-up appear. Second, the surveys need make better use of the space available to them because having to scroll from side to side makes the survey lose legitimacy and look unprofessional.

The third and fourth criteria examined in evaluating the logic puzzle’s usability are task performance and amount of learning. From the start of the tutorial to the end, users reported that they improved their ability to complete a logic puzzle. The biggest contributor to this improvement was the first stage, which explained each of the parts of a logic puzzle problem and provided a systematic way with which to examine a logic question. By the end of the tutorial, users reported that they were able to properly use the knowledge learned.

2.5 Tactical Combat Casualty Care Game-Based Tutorial

The first game-based tutorial scenario teaches the basic methods of TC3 and then guides the user in a practice simulation to apply that knowledge and develop skills. Initially a survey asks for background information and then assesses mood and prior knowledge of TC3. However, the information presented there is difficult to read. All of the questions are crammed into a sidebar while three-quarters of the screen remains blank. To make this process quicker, the questions could be made easier to read if perhaps all of them are adjusted across the entire screen and enlarged.

The overall presentation of the slideshow was clear and easy to read and understand. The information was presented in very simple text and the audio helped the user follow along. An improvement would be to make the audio more controllable, presenting more videos, and enhancing the font and photos. To make the audio more controllable there should be a button that
allows the option to turn it on or off and replay the audio if needed. This will allow users to feel more in control and able repeat the audio information if they would like to hear something again. Increasing the use of videos will enhance learning by demonstrating successful methods of TC3.

The photographs that displayed required user actions are very useful and real-life applicable. However, it would be easier for users to follow along if they were shown how to perform all the steps at one time in a video. For example, when showing how to perform a carry, it would be best to show a real-life example of how that is done so people can see it. This would help fix learning in the user’s mind and support development of mental models.

In another example, when describing how a bullet penetrates the lungs, it would be useful to have an animation with audio and text, allowing the user to follow along more easily. Fonts could also be enhanced (italicized or in bold) to highlight the importance of directions. For example, when using the phrase “Do not”, it is much more demanding and imperative if those letters are either highlighted or in bold. It would also be useful to allow users to zoom in on photos. For example, if a user wanted to see how exactly the hands should be positioned to perform a carry, it would be much easier if the view could be zoomed in rather than just telling users how to perform a carry.

After examining enhancements to the instructional content in the PowerPoint slides, there is very little to enhance in the game. It was very clear how the game controls worked within the tutorial. Very little extraneous cognitive load was imposed by the game. This ensured that the learner stayed focused on the task at hand and was not distracted by information that was not germane to the task.

One improvement to the game would be to add a narrative about what events occurred prior to the start of the game. The game is simply rushed to the action, and very little is given to the user about what to expect or what they are to accomplish during the scenario. A second area for usability improvement is to provide some immediate negative feedback for some egregious errors that occur. For example, if a tourniquet is put on a sucking chest wound, there should be immediate feedback. This could be provided either through another character in the game or through the GIFT TUI. Overall, however, this tutorial implemented very effective teaching methods and was a good source for practice.

### 3. Future Directions

While instruction provided by GIFT-based tutors is pedagogically sound and founded on validated learning principles, Holden (2013) identified several areas of learner modeling research that might improve the overall learning experience: motivation, disengagement, meta-cognition, self-regulated learning, open-learner modeling, group and collaborative learner modeling, and long-term learner modeling. To this list we have added human-tutor interaction, the application
of cognitive theories to tutoring, and the importance of narrative in tutoring as discussed in the following subsections. Each of these areas influences how users approach each training experience.

3.1 Human Interaction

A large part of learner interaction with GIFT is reading/assessing information presented on a computer screen. Reports of information overlapping or being hidden by navigation controls limit the usability and effectiveness of GIFT-based tutors. To reduce these conflicts, ARL is proposing a set of standard browser controls (like standard digital video recorder controls) placed at the bottom of the browser window (instead of in the content window) to support user navigation through the content (e.g., PowerPoint slide presentations). This will allow authors to design content that will present clearly in a browser window regardless of device (e.g., computer, smartphone).

USMA cadets identified a need for voice interaction to open up a new mode of information flow in GIFT. GIFT has been focused on providing text feedback except within game-based tutoring environments.

3.2 Application of Cognitive Theories in Tutoring

Research has been conducted by ARL (Goldberg 2013) to ascertain the modality of feedback in adaptive tutoring environments and determined that an embodied pedagogical agent (EPA) situated in GIFT’s TUI was as effective as embedding the agent directly in the simulation/game environment. Participants assigned to an EPA condition were found to perform significantly better on transfer assessments compared with subjects assigned to the audio-alone condition (e.g., so-called “Voice of God”).

So while voice interaction may be important as factor in managing cognitive load, the source of the voice seemed to be insignificant. This is good news for ITSs and virtual human developers. They can maintain a separate voice modality outside the training environment (easier/cheaper to implement than injecting voice into a virtual character in the simulation) and know that it does not add cognitive load.

Sottilare and Goldberg (2012) examined the implications of cognitive theories in ITS design and their relevance to accelerated learning and retention. The theories examined included cognitive load theory, cognitive flexibility theory, and cognitive transformation theory.

3.3 Narrative

While it may not be practical to provide every detail needed by the learner to understand instructional context, it is necessary to explain to the learner what preceded the onset of the scenario under which they will soon receive instruction. This can be a time saver. Narrative or storytelling is one way to accomplish a leveling of knowledge in a relatively short period of time.
4. References


Goldberg B. Explicit feedback within game-based training: examining the influence of source modality effects on interaction [dissertation]. [Orlando (FL)]:University of Central Florida; 2013.


Sottilare RA, Brawner KW, Goldberg BS, Holden HK. The generalized intelligent framework for tutoring (GIFT). Orlando (FL): Army Research Laboratory (US); 2012; concept paper.


INTENTIONALLY LEFT BLANK.