ARMY RESEARCH LABORATORY

An Evaluation of the Generalized Intelligent Framework for Tutoring (GIFT) from an Author's Perspective

by Robert A Sottilare, Keith W Brawner, Carson Giammaria, Katherine Mucke, Alyssa Sohn, and Michael D Matthews

ARL-SR-0303 **December 2014**

NOTICES

Disclaimers

Army Research Laboratory
Aberdeen Proving Ground, MD 21005-5425

ARL-SR-0303 **December 2014**

An Evaluation of the Generalized Intelligent Framework for Tutoring (GIFT) from an Author's Perspective

Robert A Sottilare and Keith W Brawner
Human Research and Engineering Directorate, ARL

Carson Giammaria, Katherine Mucke, Alyssa Sohn, and Michael D Matthews
United States Military Academy

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
December 2014

2. REPORT TYPE
Final

3. DATES COVERED (From - To)
20 January–4 February 2014

4. TITLE AND SUBTITLE
An Evaluation of the Generalized Intelligent Framework for Tutoring (GIFT) from an Author's Perspective

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Robert A Sottilare, Keith W Brawner, Carson Giammaria, Katherine Mucke, Alyssa Sohn, and Michael D Matthews

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
US Army Research Laboratory
ATTN: RDRL-HRT-T
Orlando, FL 32826

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-SR-0303

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
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-moderated 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 report describes the outcomes of an evaluation conducted at the US Military Academy to determine initial usability of the Generalized Intelligent Framework for Tutoring (GIFT), 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. Cadets participating in this evaluation were assigned tasks related to the author'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.

15. SUBJECT TERMS
self-regulated learning methods, intelligent tutoring system, Army Learning Model, adaptive tutoring, authoring tools, usability

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Robert A Sottilare
Unclassified	Unclassified	Unclassified	UU	18	19b. TELEPHONE NUMBER (*Include area code*) (407) 208-3007

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

Contents

INTENTIONALLY LEFT BLANK.

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 last 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; Soldiers are largely responsible to manage their own learning. From a common sense point of view it may not seem practical for each Soldier to be able to manage his/her 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) are 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 regard 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 author’s perspective, which is about what people who use GIFT to construct ITSs think about their experience and the ease-of-use of GIFT in facilitating the development of effective (pedagogically correct) ITSs.

The authoring construct within GIFT includes tools and interfaces to establish ITS standards and opportunities for reuse to reduce the need for authoring as well as automate processes to support ITS development at lower costs, time, and skill levels than currently possible today. GIFT currently supports tools for developing surveys, courses/lessons, and domain knowledge (e.g., content, question and feedback libraries). Automatic generation of expert models (standards against which learners' performance are assessed), domain models (e.g., scenario generation), and narrative content are also goals for GIFT.

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

2. Evaluation of GIFT from an Author's Perspective

Human tutoring is an established practice that we have taken advantage of for years while ITSs are substantially less well known. When attempting to create a product that mirrors the actions of an expert human tutor, the developers must create the parallels so that there is trust and ease of use within the product. In many instances, developers tend to neglect an essential aspect of the process—usability. This can take many forms, some being feedback through testing or surveys, but it is a necessary procedure to make the crucial improvements. Usability feedback tests many aspects of the product, including user frustration, effectiveness, and trust in the given good.

The first intelligent tutoring system was the program SCHOLAR, developed in 1970 in the form of an education tutor (Carbonell 1970). This was a very basic form of intelligent tutoring; many have been developed since then including the Cognitive Tutor and AutoTutor. Many tutors are specific-use tutors and not transferrable to use in other domains. Other tutors are really shell tutors or frameworks from which tutors in many domains might be authored. One example of a shell tutor is GIFT, currently being developed under a research project at the ARL. This adaptive tutoring system takes into consideration multiple facets of learning: learners, designers, instructors, authors, and researchers.

Many aspects of intelligent tutoring are important when developing these systems. These include the cost of the product and both user and developer goals for the product. Some questions that were taken into consideration while working with GIFT include the usability of the architecture, processes that were difficult to perform, and any changes that should be made to improve user experience and adaptability. It is important to evaluate and understand each user and their perspective when developing a training prototype. The goal of GIFT is to provide an effective and adaptive tool that is easy for all levels of learners and authors.

With respect to authoring, the user interface is one of the most important aspects of usability. User interface should be easy to use, support organization of the author's domain knowledge, and enhance trust in developing credible and effective tutors. The goal of the authoring tool is to give a user the ability to make an intelligent tutoring system for any purpose to enhance the learner's experience without a human instructor in the loop. This study serves to evaluate the usability of the authoring tools in GIFT.

2.1 Methods

The following sections describe the participants, apparatus, procedure, and results of the authoring evaluation of GIFT.

2.2 Participants

The evaluators of GIFT authoring tool usability are "firsties" (senior-level cadets) at USMA.

2.3 Apparatus

The apparatus used in this evaluation was the intelligent tutoring system architecture GIFT. The main tools used within this evaluation were the authoring tools within GIFT.

2.4 Procedure and Results

The evaluation was conducted by working through each tool in the Administrator Tools of the GIFT Monitor Module to implement changes as requested by the ARL Adaptive Tutoring Team. We only completed changes to Condition 1 of the TC3 Training Scenario. We did not access the help menus or GIFT forums for guidance so as not to detract from the following evaluation activities.

- With the addition of the Learning Management System, there was no longer a need for the course to begin with a demographic survey. To remove the survey, we entered the Course Authoring Tool (CAT) and deleted the node associated with this choice (Explicit Feedback Demographics). In addition to removing the survey, we had to remove the guidance that prompted the survey.
- To capture the interests of users for development of a new leadership course, we first entered the Survey Authoring System (SAS), where we built questions into the question bank. We had a number of options for formatting these questions; however, we left them as free-response questions. Once we built the questions, we inserted them in to the new survey we created—Interests (this survey was limited to the Feedback Experiment folder). From there, we added the survey to the end of the scenario by creating 2 more nodes: 1) one guidance node to prompt the user for the survey and 2) the survey itself.
- Adding extra feedback when a user made the same mistake twice required use of the Domain Knowledge Editor. Under the "you are part of a unit" strategy, we added the additional feedback StayCloseNow.

- Removing the exam scenario required deleting it and its associated nodes, then creating a new XML (extensible markup language) document solely for the exam. We created this new XML with the CAT, giving it guidance before and after the exam and an after action review (AAR) at the end.
- To record self-reports, we entered the Sensor Configuration Tool and made the Self-Assessments file externally available by changing it from false to true.
- To show the answers following the pretest, we included an AAR to show the users' behaviors during the test.
- In the CAT, we copied the Training Application node containing the initial PowerPoint training and placed it after the second scenario so that it would repeat.
- We were unable to export the revised file for Condition 1 of the TC3 Training Scenario.

3. Conclusions

When discussing the usability aspects of GIFT, recognize that this perspective comes from users with virtually no experience in editing pre-existing scenario files in GIFT. Additionally, much of what was done in the procedures is not validated work, meaning we made changes but were unable to examine them for correctness. This inability to see what changes were accurate may or may not skew our perception of the program's difficulty. Aspects of our efforts are discussed in the following.

1. While working with GIFT, what part of the process was intuitive or easy to accomplish?

   After the initial difficulties of starting up GIFT, several features are intuitive to even the most inexperienced user. Expanding, moving, adding, and deleting nodes within the CAT is extremely simple because it is linked with the mouse's right-click feature. The challenge arises in understanding what role the node plays in the system and what its functions are. Within the SAS, creating questions, surveys, and survey contexts was not complicated. The survey system benefited from the tabs separating the various sections, and it was intuitive to discern what components are necessary for a survey in the first place. Still, GIFT could further streamline this system by essentially creating a "wizard" or job aid that sequentially walks authors through building a survey. An example of this system would be 1) receive a prompt to create a new survey, 2) be asked to add questions to it (with the option of choosing from a library of already created questions), and 3) when the survey is done, users would have the option to limit what contexts the survey could be use in.

2. What part of the process was difficult?

   From the beginning, initializing the program required authors to access a sequence of scripts within the main folder. A single program that would open all necessary scripts would mitigate this obstacle. Once the module was open and all systems were running, it was not apparent which section was for authoring without additional guidance. There needs to be consistency in terminology; this is something discussed in later questions.

   When inside the administrative tools, there was a list of tools that lacked any description beyond their title, leading us to explore each option. Fortunately the instructions' hints pointed us in the right direction. After loading the experimental files, again we were forced to navigate through countless nameless hierarchies. Instead of just labeling these sections as to what type of node they are, they should be linked to an author-input name (i.e., Introduction, Demographic Survey, Stage 1). These ambiguous hierarchies plague most of the system, and ambiguous labels for different sections only worsen this problem. As authors we lacked even a basic understanding of what Domains, Sensors, and other important options were. The lesson learned here is to read the authoring documentation or to implement a tutor (e.g., GIFT tutors GIFT) or job aid to guide the authoring process.

   Ultimately we were unable to export the finished product. Errors are expected but the troubleshooting to resolve such errors should not simply point authors to logs of cryptic computer language (unless this level of programming skill is a specified prerequisite). GIFT is definitely not a pick-up-and-use system at this point.

3. What prerequisite skills are needed to use GIFT as it is today?

   To effectively use GIFT there are 3 basic prerequisites: 1) a basic understanding of computers (mostly assumed), 2) an introduction to the terminology (the labeling in GIFT does not automatically elicit its meaning), and 3) a summary of the logic that links tools together (how the different tools interrelate CAT, the Domain knowledge file Authoring Tool [DAT], SAS, etc.).

4. What changes would you make to GIFT to improve the experience from your assigned perspective? Give specific examples of how you would make it more usable.

   The authoring experience would greatly benefit from editing a scenario or system in a single location; this difference could reduce much of the confusion involved in handling GIFT. The tools used for different aspects (CAT, DAT, SAS, etc.) could then be accessed from a toolbar-type menu. Microsoft Word is an excellent example of this. In Word authors can always see the end product and easily navigate throughout it, and when they need to access a tool to change something within it they refer to the toolbar, thus users can see how their changes affect the overall product.

Another example, more closely linked to what GIFT is capable of, is any sort of HTML (hyper text markup language) editor. These programs allow for quick previewing, navigation through multiple hierarchies, and still do not sacrifice the complexity necessary to create elaborate systems.

5. What changes would you recommend to make GIFT more adaptive and able to support tutoring without a human tutor?

   The option of having “novice” and “advanced” versions would be extremely helpful in limiting what more advanced options are available to authors, yet it is not entirely feasible to build 2 versions of GIFT. Instead, GIFT could have built-in tips (i.e., hovering over highlight sections opens a pop-up description). These tips would not interfere with familiar users but would allow users to learn about the system without disengaging or referring to external resources. This primary change would make the exploration much more practicable.

4. References

Bloom BS. The 2 sigma problem: the search for methods of group instruction as effective as one-to-one tutoring. Educational Researcher. 1984;13(6):4–16.

Carbonell. AI in CAI: an artificial intelligence approach to computer-assisted instruction. IEEE Transactions on Man-Machine Systems. 1970;11:190–202.

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

VanLehn K. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. Educational Psychologist. 2011;46(4):197–221.

1 (PDF) DEFENSE TECHNICAL
INFORMATION CTR
DTIC OCA

2 (PDF) DIRECTOR
US ARMY RESEARCH LAB
RDRL CIO LL
IMAL HRA MAIL & RECORDS MGMT

1 (PDF) GOVT PRINTG OFC
A MALHOTRA

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRM D
T DAVIS
BLDG 5400 RM C242
REDSTONE ARSENAL AL 35898-7290

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRS EA DR V J RICE
BLDG 4011 RM 217
1750 GREELEY RD
FORT SAM HOUSTON TX 78234-5002

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRM DG J RUBINSTEIN
BLDG 333
PICATINNY ARSENAL NJ 07806-5000

1 (PDF) ARMY RSCH LABORATORY – HRED
ARMC FIELD ELEMENT
RDRL HRM CH C BURNS
THIRD AVE BLDG 1467B RM 336
FORT KNOX KY 40121

1 (PDF) ARMY RSCH LABORATORY – HRED
AWC FIELD ELEMENT
RDRL HRM DJ D DURBIN
BLDG 4506 (DCD) RM 107
FORT RUCKER AL 36362-5000

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRM CK J REINHART
10125 KINGMAN RD BLDG 317
FORT BELVOIR VA 22060-5828

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRM AY M BARNES
2520 HEALY AVE
STE 1172 BLDG 51005
FORT HUACHUCA AZ 85613-7069

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRM AP D UNGVARSKY
POPE HALL BLDG 470
BCBL 806 HARRISON DR
FORT LEAVENWORTH KS 66027-2302

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRM AT J CHEN
12423 RESEARCH PKWY
ORLANDO FL 32826-3276

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRM AT C KORTENHAUS
12350 RESEARCH PKWY
ORLANDO FL 32826-3276

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRM CU B LUTAS-SPENCER
6501 E 11 MILE RD MS 284
BLDG 200A 2ND FL RM 2104
WARREN MI 48397-5000

1 (PDF) ARMY RSCH LABORATORY – HRED
FIRES CTR OF EXCELLENCE
FIELD ELEMENT
RDRL HRM AF C HERNANDEZ
3040 NW AUSTIN RD RM 221
FORT SILL OK 73503-9043

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRM AV W CULBERTSON
91012 STATION AVE
FORT HOOD TX 76544-5073

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRM DE A MARES
1733 PLEASONTON RD BOX 3
FORT BLISS TX 79916-6816

8 (PDF) ARMY RSCH LABORATORY – HRED
SIMULATION & TRAINING
TECHNOLOGY CENTER
RDRL HRT COL G LAASE
RDRL HRT I MARTINEZ
RDRL HRT T R SOTTILARE
RDRL HRT B N FINKELSTEIN
RDRL HRT G A RODRIGUEZ
RDRL HRT I J HART
RDRL HRT M C METEVIER
RDRL HRT S B PETTIT
12423 RESEARCH PARKWAY
ORLANDO FL 32826

1 (PDF) ARMY RSCH LABORATORY – HRED
HQ USASOC
RDRL HRM CN R SPENCER
BLDG E2929 DESERT STORM DRIVE
FORT BRAGG NC 28310

1 (PDF) ARMY G1
DAPE MR B KNAPP
300 ARMY PENTAGON RM 2C489
WASHINGTON DC 20310-0300

12 (PDF) DIR USARL
RDRL HR
L ALLENDER
P FRANASZCZUK
J LOCKETT
RDRL HRM
P SAVAGE-KNEPSHIELD
RDRL HRM AL
C PAULILLO
RDRL HRM B
J GRYNOVICKI
RDRL HRM C
L GARRETT
RDRL HRS
D HEADLEY
RDRL HRS B
M LAFIANDRA
RDRL HRS C
K MCDOWELL
RDRL HRS D
A SCHARINE
RDRL HRT T
R SOTTILARE

INTENTIONALLY LEFT BLANK.