

The 2017 Overview of the GIFT Authoring Experience

Scott Ososky

Oak Ridge Affiliated Universities
U.S. Army Research Laboratory

INTRODUCTION

One of the primary goals of the Generalized Intelligent Framework for Tutoring (GIFT) is to reduce the time and skill required to create adaptive training. Achieving those goals within the GIFT platform is enabled by a set of authoring tools and associated resources, known collectively as the Authoring Experience. The current paper will discuss the state of the GIFT Authoring Experience, as well as how feedback and data from the community is informing user-centered design efforts within the GIFT authoring tools.

The discussion will focus specifically on work conducted over the past year, including improvements to UI and workflows, as well as additional features that have been added to the authoring tools. Those enhancements will be described in the context of field research, conversations with the user community, and case studies of GIFT use in real-world settings. Finally, this paper will preview aspects of the GIFT authoring experience that are under investigation, and the authoring tool enhancements that are intended to result from ongoing user-centered design research efforts. This paper is intended to benefit both new and experienced GIFT users, and may be of interest to anyone conducting design research or developing interfaces and user experiences for computer-based productivity tools.

NEW AUTHORING TOOLS

The new authoring experience that was described in last year's GIFT User's Symposium proceedings is now available (Ososky, 2016a; Ososky & Brawner, 2016) at GIFT Cloud (cloud.tutoring.org). This new experience is centered providing a unified interface, making language within the authoring tools more consistent and intuitive, and building user-centered tools that support authors' goals. Readers interested in the history and evolution of the GIFT authoring tools are encouraged to read the references contained within this paragraph (Ososky, 2016b; Ososky & Sottolare, 2016).

Specifically, the new authoring experience revolves around a visual course building interface within the Course Creator. From within this interface, all other core aspects of course authoring are accessible to the user. The course flow timeline interface was redesigned based on a flow chart (or discrete event process) metaphor with simple drag-and-drop functionality. The visual structure of the course more accurately suggests the sequencing functions that are available to course authors. The design intent was to evoke a *mental model* of similar, more familiar interfaces in order to make this authoring task more intuitive for novice users (Figure 1). Available course objects are displayed in the toolbox on the left-hand side of the interface. Authors can drag and drop objects onto the timeline in any position. Objects already on the timeline can be re-ordered or deleted as needed.

Some course objects have been renamed to provide a better indication of their functionality and/or breadth, respectively. Complementary to the course object interface is the existence of an on-demand help window that appears in the lower-right hand corner of the interface. Currently, interacting with any of the course objects within the interface displays information about that object within the window. This is useful for new authors, as well as more experienced authors who are trying to decide between multiple viable course objects to complete their course sequencing goals.

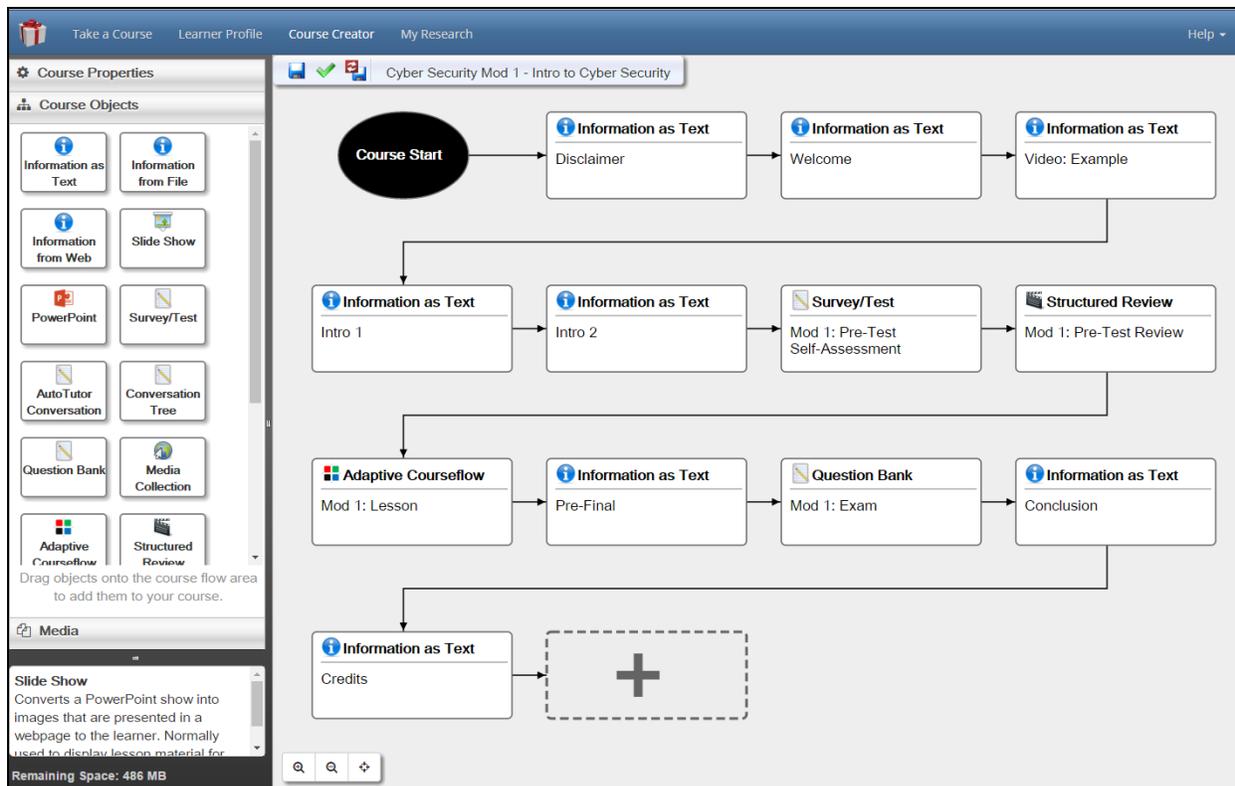


Figure 1: Visual course flow editor uses familiar interaction metaphors to allow authors to quickly create course outlines.

The efficient editing of course objects is also central to the new authoring experience. Recognizing that authoring an adaptive tutor required complex yet usable authoring tools, our design approach to efficiency was inspired by the similarly complex design video game development tools (Lightbown, 2015). The new authoring experience adopts a design philosophy that reduces the number of pop-ups on screen, and keeping multiple pieces of information in view in order to reduce the memory/recall burden on the author. To that end, editing a course object is as simple as clicking it on the timeline. Now, instead of a pop-up, the course object editor opens in a side frame on the right-hand side of the workspace (Figure 2).

Each course object has a different set of editing and configuration options; some editors require more screen space than others in order to be in full view of the author. Therefore, each of the three primary *panels* of the Course Creator are resizable. The course object editing UI can also toggle into *full screen* mode as the authoring work shifts from sequencing to configuring course objects. The editing area, by default, displays the most recent course object that the author has clicked in the timeline. Individual course object editors can be *pinned* to the editing frame and quickly accessed as a series of tabs (like a web browser) should the author need to view or edit multiple course objects simultaneously.

The object that is currently in-focus within the editing frame is also highlighted on the timeline with an animated blue-dashed outline around the object. The intent of that design is to help the author make the connection between what is being edited and where that object exists in the timeline. This is useful when courses have multiple objects that appear similar to one another on the timeline, such as an informational message, or a survey / test (Figure 2).

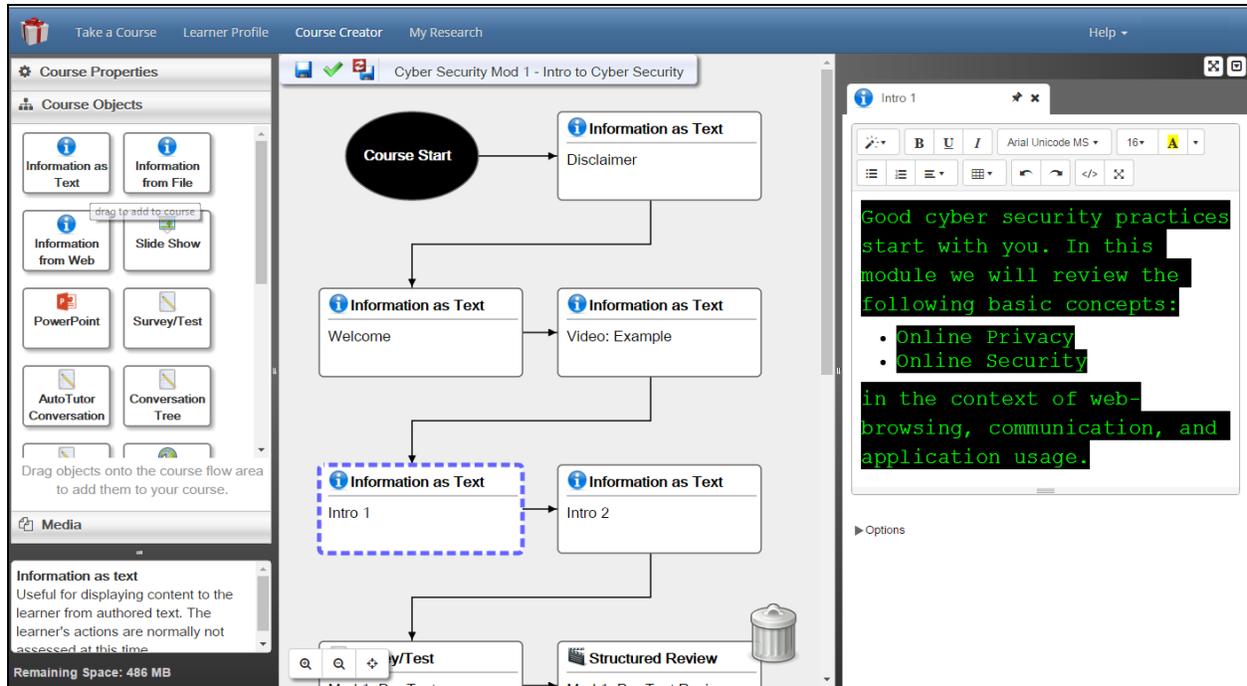


Figure 2: The editing frame appears when interacting with objects on the timeline. All three sections of the Course Creator interface are resizable, based on user needs and preferences.

The survey authoring interface has also been completely re-designed. Specifically, we focused on reducing users' physical effort within the system, or *excise*, by reducing the number of mouse clicks and keyboard commands required to build an individual question. Further, system-level configuration parameters, such as the collection of surveys for a particular course (i.e., survey context) are now automatically managed by the GIFT and are invisible to the author.

The survey authoring interface is now presented in a WYSIWYG style format (Figure 3). After writing the question text, and a first response item, additional response items are automatically added so that the author can quickly complete the question, and then easily get a sense of how that question will look when the learner sees it. Text fields within the survey composer also use the same, familiar rich-text editor that is found within the Information Course Object types. The survey composer now supports quickly copying and moving questions within the interface, further increasing authoring efficiency. Configuration options are organized within different parts of the interface in order to maintain authoring efficiency and the visual integrity of the survey layout. Options such as "force response" or "multi-select" now appear in a side frame, which is dynamically updated as individual questions are selected. The survey system also supports multi-select for batch operations on groups of questions.

Scoring mode (Figure 3, top-middle) is a new feature of the survey composer intended to reduce authors' cognitive workload through progressive disclosure of information. The new mode functions as a toggle between *writing* surveys and *scoring* surveys. Activating scoring mode temporarily locks the questions for editing; the UI adds a series of score boxes to each response for a particular question (Figure 4). Further, if the survey is linked to a set of learning concepts, additional options will appear on each question that will allow the author to associate the question with a concept(s) and set a difficulty for that question. Toggling between the Writing Mode and Scoring Mode reduces visual clutter, and allows authors to focus on one specific aspect of survey composition. The benefit of the two modes is most apparent when questions have complex scoring mechanisms, such as a matrix of responses.

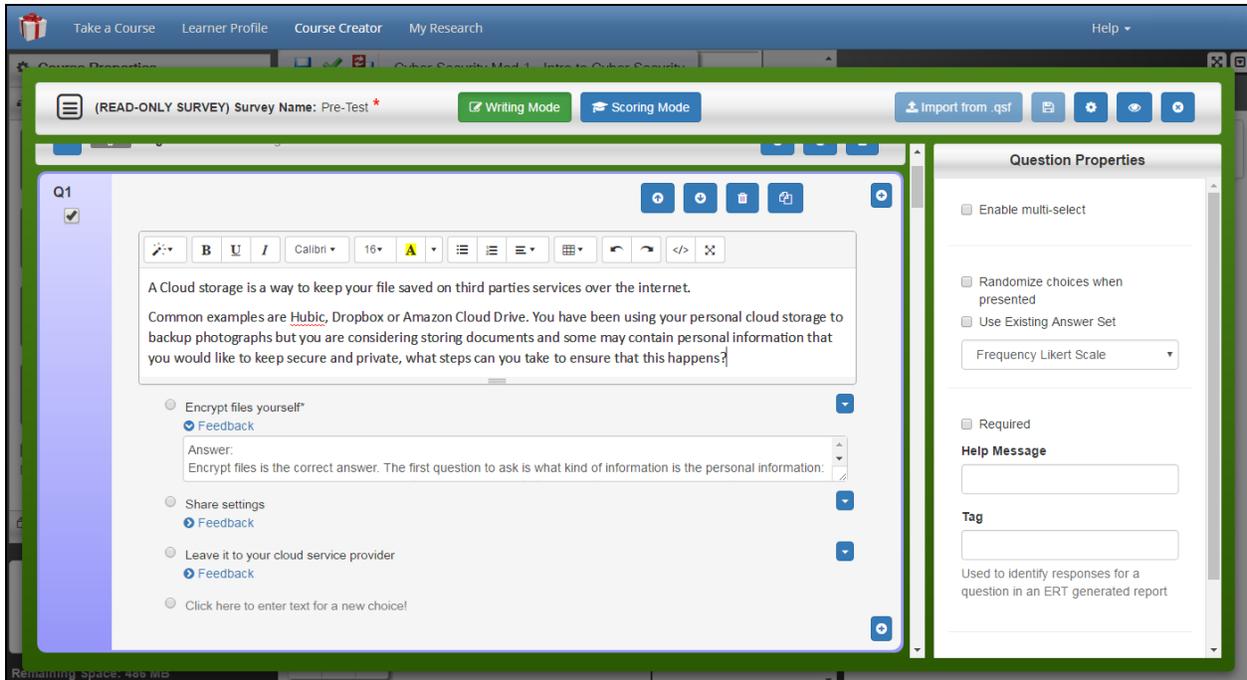


Figure 3: The survey composer has been completely redesigned as a WYSIWYG editor with rapid survey creation in mind. Configuration options appear on the right-side frame. Scoring options are enabled through a separate scoring mode toggle at the top of the interface.

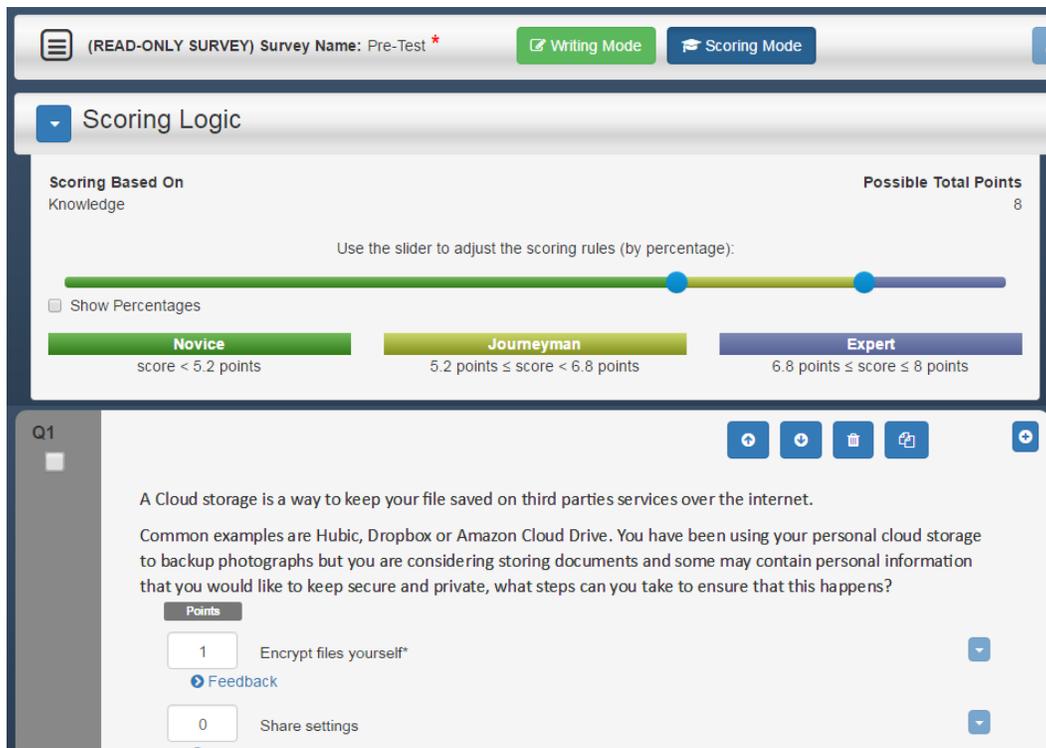


Figure 4: Survey with scoring mode activated. Scoring logic dynamically updates with score values.

Scoring logic also appears at the top of each scored survey. The design intent was to reduce the possibility of making errors in scoring a survey, by selecting UI controls that were appropriate to the task. As authors set the numbers of possible points for each question, the scoring logic automatically calculates the maximum total points. This maximum is used to set a slider for specifying the scoring logic to sort scores into three bins: *novice*, *journeyman*, and *expert*. The range of possible scoring outcomes is clearly shown on screen under each bin. Scoring can be set based on a percentage score, or raw point totals (scoring on fractional points is also possible using the sliders). The author can also set the two sliding points on top of one another to create a binary classification system, if desired (e.g., pass / fail).

Lastly, an “import from .qsf” feature is now available in the survey composer. This is an example of a function that was added in direct response to community feedback. It was recognized that some GIFT users already had experience in creating survey content using Qualtrics, a survey creation and data collection tool, used primarily for marketing research. Qualtrics exports content from their system in a .qsf format. The “import from .qsf” feature in GIFT scans the export file for survey questions and takes one of three actions: imports the question as it appeared in Qualtrics, adapts a Qualtrics question type to a GIFT question type, skips question import and notifies the author. Because the primary purposes of GIFT and Qualtrics are different, the question types supported by each are not a one-to-one mapping. Therefore, GIFT imports questions from “.qsf” files when the integrity of the question can be maintained.

Additional enhancements to the authoring tools are in various stages of development and release. We continue to collect feedback from current and potential authors in order to determine the suitability and efficiency of the tools in order to provide a positive and productive experience for authors of all skill levels. We encourage readers to examine the authoring tools further by visiting cloud.gifttutoring.org.

USER TESTING

The new user experience of the GIFT Authoring tools was largely inspired by user feedback and comparative analysis; however, there is currently little formal data available to measure and evaluate the usability, efficiency, and perception of current and future authoring experiences within GIFT Cloud. To that end, a round of formal usability evaluation was recently completed, regarding elements of the new authoring tools, described above. The objectives of this research were to evaluate the usability of the current version of the GIFT Authoring Tools (GAT) by: (1) establishing a baseline of task-performance measures, (2) establish a baseline for the overall usability of GIFT, and (3) identify potential design concerns to be addressed in order to improve the efficiency, productivity, and satisfaction with GIFT.

This section reports on a subset of the data collected during the evaluation, with a comprehensive report in preparation for publication in the coming months. Specifically, this section will describe some of the subjective, quantitative survey data collected throughout the study.

Participants

The study was conducted with twelve participants that were new to GIFT, are instructors and/or instructional designers and are familiar with military training / instruction. As this was conducted as a usability evaluation of the GIFT authoring tools, no additional demographic data was collected about the participants.

Procedure and Tasks

Participants interacted with GIFT on a laptop computer and were asked to perform eight authoring tasks (Table 1). These tasks were representative of typical authoring tasks within the GAT and were specifically designed around elements of the interface that were recently redesigned (e.g., course editor, survey system). Prior to the usability tasks, participants received training in the form of a preparatory overview of adaptive training, the GIFT platform, how GIFT enables adaptive training, and superficial information regarding GIFT’s primary features and functions. Training was administered by PowerPoint and in-person lecture.

Table 1: Description of usability tasks and scenario descriptions

Task description (not given to participant)	Scenario (given to participant)
1. Add a new “Information as Text” course object to the course. Name it, “Welcome”.	1. You have been tasked with creating some Cyber Security training content. You want to use GIFT to create the a basic lesson. First, you want to display a welcome message that will be displayed to the learner when they start the course. How would you go about setting up the first screen a learner would see in a training lesson?
2. Edit the content of the “Welcome” course object. Make the text bold and blue.	2. You now want to edit the text of the Welcome screen to read, “Welcome to the course”. You want text to be bolded and blue. How would you go about editing the text in GIFT?
3. Add two course concepts to the Course, “Internet Privacy” and “General Security”.	3. The lesson you are creating will cover two concepts, Internet Privacy and General Security. How would you go about specifying these lesson concepts inside of GIFT?
4. Add a Media course object to the course, add an image from the desktop to the course object.	4. As part of the lesson materials, you want to display an image to the learner regarding protecting accounts with two-factor authentication. You have a copy of the relevant image on your computer desktop. How would you go about adding this image to your lesson?
5. Add a three item questionnaire to the course, name it “Pre-Test”.	5. You want to add a simple pre-test survey to the lesson in order to gauge learners’ existing knowledge. You have 3 sample questions contained within a notepad file on the computer. How would you go about administering the survey within GIFT?
6. Add a Slide Show course object to GIFT, use a PowerPoint show file from the desktop.	6. Some of your existing lesson material includes a slide show in the form of a PowerPoint show. You want to display this content to the learner after the pretest. How would you go about adding this content to GIFT in the desired order?
7. Edit the existing Adaptive Courseflow course object, by adding two web links to the Rules and Examples quadrants, respectively. Tag the content with appropriate metadata.	7. Your Cyber Security lesson now includes an adaptive component, made up of lesson materials and a quiz. You want to add some supporting web-links to the lesson material contained within the adaptive portion of the course. How would you go about adding this content in GIFT, and making sure it is tied to the correct lesson concepts?
8. Make a copy of the completed course.	8. You have recently completed the first lesson of the Cyber Security course, and are ready to get started on the second lesson. Instead of starting from scratch, you want to use the first lesson as a template. How would you go about making a copy of the first completed lesson in GIFT?

After the training, participants were presented with the list of tasks they would attempt to perform, and were asked to rate their *expectation* of how easy or difficult each task would be (using a 7-point scale). Expectation / experience ratings (Albert & Dixon, 2003) leverage a single-item questionnaire which is administered before and after a usability task, respectively:

- Before all tasks (expectation rating): “How difficult or easy to you expect this task to be?”
- After each task (experience rating): “How difficult or easy did you find this task to be?”

Those questions appeared with a seven point response scale where (1) is Very Difficult and (7) corresponds with Very Easy. The post-task ratings have been reported to be significantly correlated ($r = 0.46$, $n = 227$, $p < 0.0001$) with objective task completion rates and times (Tedesco & Tullis, 2006). Differences between expectations and experience ratings can also be used to identify opportunities within a system.

Following the completion of all usability tasks, participants were asked to complete the System Usability Scale (SUS), is an industry standard tool for quickly measuring the usability of systems and software (Brooke, 1986, 1996). The survey was designed to be administered after all tasks have been completed, but before any additional discussion or debriefing. The survey consists of 10 questions, with a 5-point response scale where 1 = strongly disagree, and 5 = strongly agree.

Preliminary analyses

For each of the eight usability tasks, the expectation and experience ratings, respectively, were averaged across all participants. The resulting pairs of values allow us to quickly gauge the health of each of the tasks, by plotting pairs of values along an X-Y axis (Figure 5). The axis is segmented into four quadrants, representing opportunities for improvement (i.e., perceived to be easy, but was difficult), features that could be promoted (i.e., perceived to be difficult, but was easy), and so on (Tullis & Albert, 2013, p. 132). For this particular usability evaluation, each of the eight evaluation tasks were, on average, expected to be relatively easy and were perceived as relatively easy once completed.

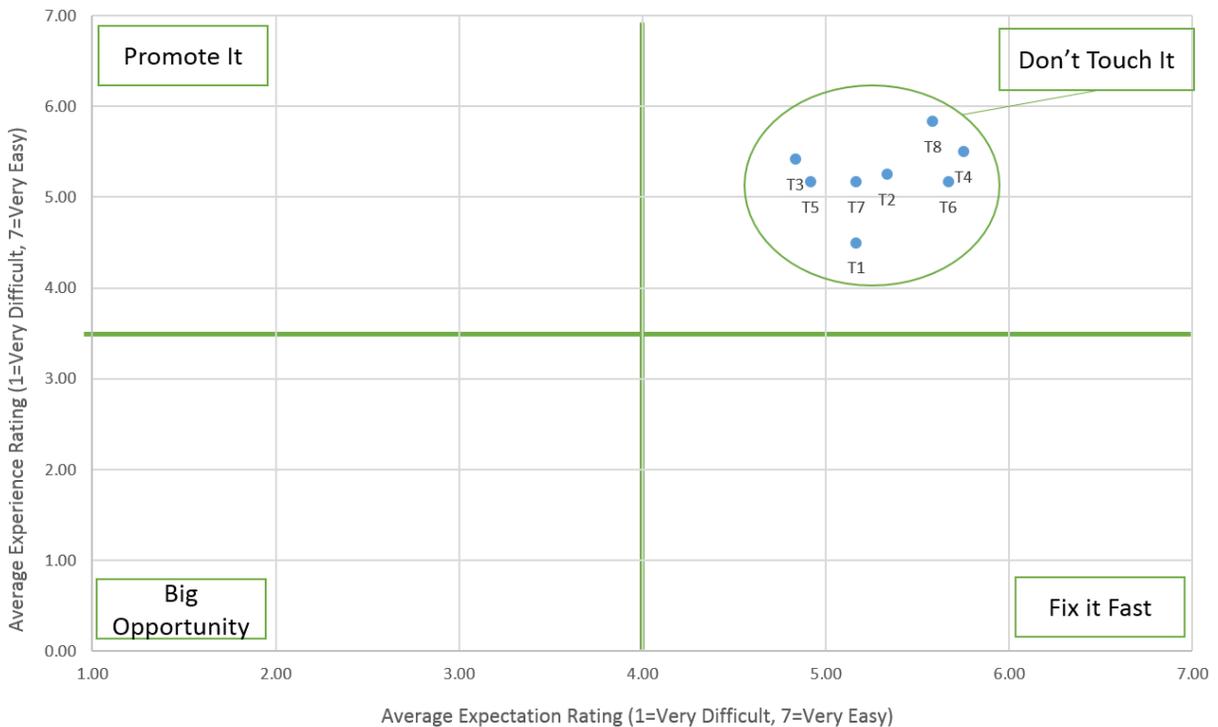


Figure 5: Average subjective expectation and experience ratings by task

The expectation and experience ratings provide a subjective snapshot of a system at the individual task level. The SUS, by comparison, provides a snapshot of the overall subjective usability of a system. The SUS score is calculated by reverse coding the even numbered questions, summing the data from all questions, and then adjusting the scale of the score to 100, however that is not a percentage score. Grades

are assigned to scores based on the percentile into which the score falls (like a bell-curve). The GAT currently does not have its own SUS reference score; however, an external reference database of other system SUS scores is available for comparison (Sauro & Lewis, 2012, pp. 198-200). For the current usability study, the GIFT authoring tools yielded an average SUS score of 58.75, which ranks in the 29th percentile of all other systems in the database (roughly a D or D+ ranking). While there is room for improvement within the GAT, that result represents a modest improvement over a previous usability survey in which *perceived ease of use* for GIFT was measured (Holden & Alexander, 2015).

Preliminary discussion

The current study offers a snapshot in time regarding the subjective usability of GIFT at the system and task levels, respectively. The results reported in this paper indicates that GIFT has experienced a modest usability gain from the last usability survey, but there is room for improvement. Interestingly, subjective experiences measured at the task level were positive, which is encouraging because these were the areas of GIFT which received the most design attention. It is expected that the overall usability of the system will continue to improve as other interfaces within the GAT receive targeted design attention.

Further, there is additional objective performance (i.e., types of errors) and subjective qualitative data (i.e., task times, error rates) that is currently under analysis. There will be more conclusions to be drawn from the study, once those data have been analyzed, and opportunities for improvement can be triangulated from multiple data streams. For instance, observation of and discussion with participants identified media upload and management as an early candidate for additional design attention in the coming year. Finally, this research will also now serve as a new baseline for internal comparison with future versions of the GAT, where *expectation ratings* and the SUS are used, respectively.

LOOKING AHEAD

The new interfaces described in this paper provide usable tools and experiences that are intended to help authors configure tutors, sequence courses, and integrate instructional content into those courses. In theory, it would be desirable to have specific content relative to each adaptive component of a tutor, with respect to features within a learner model. However, *creating* instructional content suitable for adaptive tutors is still a resource intensive task, even when existing source material is available. In the coming year, the GIFT team will investigate forward-thinking solutions that will support authors in creating tailored learner experiences within a tutor, by reducing the time and effort spent in content preparation.

GIFT can already monitor and interact with external simulation and serious game systems. Technologies are being investigated that would allow single training or game scenario to be automatically reconfigured over hundreds of variations to provide suitable adaptations within GIFT. This is known as automated scenario generation (Zook et al., 2012) or evolutionary scenario generation (Luo, Yin, Cai, Zhong, & Lees, 2016). Scenario permutations would be automatically ranked by the system based on an author's specified learning objectives. Further filtering can be done by the author and their human collaborators. This would allow a GIFT tutor to respond in real-time to a learner's needs without minimal additional development. This capability would also benefit learners in the generation of replayable, simulation-based practice material within a course. For more information see (Sottolare & Brawner, 2017).

Further, adaptation within a tutor is commonly operationalized based on content: a remedial video, targeted feedback, more difficult quiz questions, and so on. Alternatively, is it thought that there are ways to provide an adaptive tutoring experience to learners that is, in-part, content independent. For instance, inspiration can be drawn from gamification, or the application of game elements to non-gaming contexts. Gamification is typically implemented as a one-size-fits-all salutation, but adaptive systems afford the

ability to intelligently tailor gamification-type features within the tutor-user interface that harmonize with a learner's motivation, grit, etc. (Ososky, 2015). Personalization is another way in which content-independent adaptive tutoring can be achieved. The design of personalization varies by implementation, though the premise is consistent. Personalizing learning content, with a learner's preferences or other information from their profile, can yield positive benefits on engagement and knowledge retention (A. Sinatra, 2016; A. M. Sinatra, 2015).

CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

The GIFT team remains committed to providing an authoring experience for GIFT which results in greater efficiency with which tutors can be created and managed. This paper described an effort to redesign significant portions of the GAT in service of those goals. Preliminary results from a user test support the notion that the revised authoring tools were met favorably; however, there remain opportunities to improve the authoring experience as a whole. GIFT should continue to provide design attention to authoring interfaces that are found to be difficult to use and/or understand, such as media management and real-time assessment authoring.

Interfaces, however, are only part of a user's experience with a system. Future research and design effort should continue to work toward creating supporting material such as authoring *wizards*, tutorial videos, and demonstration courses that showcase the benefit of the GIFT platform (i.e., *killer apps*). In support of a comprehensive user experience, a GIFT "Summer Camp" is currently in development, which will offer in-person training directly to members of the GIFT community. While the primary goal of this offering is to support new and current authors, the GIFT team will also likely find much to learn about how individuals use GIFT as well.

Finally, the GIFT project should continue to peruse opportunities to regularly measure the usability and efficiency of the GAT. That research needs to be able to keep pace with that of software development to enable both efforts to inform one another. Future research should also leverage results from the current user study as a point of comparison in order to measure improvement, alongside the use of publically available reference databases as a general indication of usability. In doing so, data-informed decisions can be made regarding the allocation of design and development effort for the GIFT authoring experience.

As always, we encourage you, the reader, to *join the conversation* at GIFTTutoring.org. The members of the GIFT community have a valuable opportunity to help shape how features are designed and implemented into GIFT. The GIFT development team encourages members of the GIFT community to continue to communicate feedback, issues, suggestions, and results (of research) in order to help us provide the useful tools, powerful technologies, and positive user experiences that will make adaptive tutoring technology accessible and valuable to the broadest possible audience.

REFERENCES

- Albert, W., & Dixon, E. (2003). Is this what you expected? The use of expectation measures in usability testing. *Proceedings of the Usability Professionals Association 2003 Conference, Scottsdale, AZ.*
- Brooke, J. (1986). System usability scale (SUS): a quick-and-dirty method of system evaluation user information. *Reading, UK: Digital Equipment Co Ltd.*
- Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry, 189(194), 4-7.*
- Holden, H., & Alexander, M. (2015). An analysis of GIFT user experiences and design recommendations for enhancing GIFT's usability. In R. Sottolare (Ed.), *Proceedings of the Second Annual GIFT Users Symposium* (pp. 83-89). Orlando, FL: U.S. Army Research Laboratory.
- Lightbown, D. (2015). *Designing the User Experience of Game Development Tools*. Boca Raton, FL: CRC Press.

- Luo, L., Yin, H., Cai, W., Zhong, J., & Lees, M. (2016). Design and evaluation of a data-driven scenario generation framework for game-based training. *IEEE Transactions on Computational Intelligence and AI in Games*.
- Ososky, S. (2015). Opportunities and Risks for Game-Inspired Design of Adaptive Instructional Systems. In D. D. Schmorrow & M. C. Fidopiastis (Eds.), *Foundations of Augmented Cognition: 9th International Conference, AC 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015, Proceedings* (pp. 640-651): Springer International Publishing.
- Ososky, S. (2016a). Designing the user experience of the GIFT Cloud authoring tools. In R. Sottilare & S. Ososky (Eds.), *Proceedings of the 4th Annual GIFT Users Symposium* (Vol. 4, pp. 145-156). Orlando, FL: U.S. Army Research Laboratory.
- Ososky, S. (2016b). Practical requirements for ITS authoring tools from a user experience perspective. In D. D. Schmorrow & M. C. Fidopiastis (Eds.), *Foundations of Augmented Cognition: Neuroergonomics and Operational Neuroscience: 10th International Conference, AC 2016, Held as Part of HCI International 2016, Toronto, ON, Canada, July 17-22, 2016, Proceedings, Part II* (pp. 55-66): Springer International Publishing.
- Ososky, S., & Brawner, K. (2016). The GIFT 2016 community report. In R. A. Sottilare & S. Ososky (Eds.), *Proceedings of the 4th Annual GIFT Users Symposium* (Vol. 4, pp. 63-74). Orlando, FL: U.S. Army Research Laboratory.
- Ososky, S., & Sottilare, R. A. (2016). *Heuristic evaluation of GIFT authoring tools*. (ARL-SR-0353). Orlando, FL: US Army Research Laboratory Retrieved from http://www.arl.army.mil/www/default.cfm?technical_report=7645.
- Sauro, J., & Lewis, J. R. (2012). *Quantifying the user experience: Practical statistics for user research*: Elsevier.
- Sinatra, A. (2016). A Review of Self-Reference and Context Personalization in Different Computer-Based Educational Domains. In R. A. Sottilare & A. Graesser (Eds.), *Design Recommendations for Intelligent Tutoring Systems: Volume 4-Domain Modeling*. Orlando, FL: U.S. Army Research Laboratory.
- Sinatra, A. M. (2015). A personalized GIFT: recommendations for authoring personalization in the generalized intelligent framework for tutoring. *International Conference on Augmented Cognition*, 675-682.
- Sottilare, R., & Brawner, K. (2017). *Automated Scenario Generation Broad Agency Announcement*. US Army Research Laboratory.
- Tedesco, D., & Tullis, T. (2006). A comparison of methods for eliciting post-task subjective ratings in usability testing. *Usability Professionals Association (UPA), 2006*, 1-9.
- Tullis, T., & Albert, W. (2013). *Measuring the user experience: collecting, analyzing, and presenting usability metrics*. Waltham, MA: Elsevier Inc.
- Zook, A., Lee-Urban, S., Riedl, M. O., Holden, H. K., Sottilare, R. A., & Brawner, K. W. (2012). Automated scenario generation: toward tailored and optimized military training in virtual environments. *Proceedings of the international conference on the foundations of digital games*, 164-171.

ACKNOWLEDGMENTS

The research described herein has been sponsored by the U.S. Army Research Laboratory. The statements and opinions expressed in this article do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

ABOUT THE AUTHOR

Dr. Scott Ososky is a Postdoctoral Fellow within the US Army Research Laboratory's Human Research and Engineering Directorate (ARL-HRED). His current research examines mental models of adaptive tutor authoring, including user experience issues related to authoring tools and interfaces. His prior work regarding mental models of human interaction with intelligent robotic teammates has been published in the proceedings of the Human Factors and Ergonomics Society, HCI International, and SPIE Defense & Security annual meetings. Dr. Ososky received his Ph.D. and M.S. in Modeling & Simulation, as well as a B.S. in Management Information Systems, from the University of Central Florida.