Recommendations for use of Adaptive Tutoring Systems in the Classroom and in Educational Research

Anne M. Sinatra[1], Scott Ososky[1,2], Robert Sottilare[1], & Jason Moss[1]

[1]U.S. Army Research Laboratory, Orlando, FL USA
[2]Oak Ridge Associated Universities, Orlando, FL USA

anne.m.sinatra.civ@mail.mil, scott.j.ososky.ctr@mail.mil,
robert.a.sottilare.civ@mail.mil, jason.d.moss11.civ@mail.mil

Abstract. The current paper and presentation provide background on the different uses of intelligent tutoring systems (ITSs) in context of course instruction, discusses specific instructor considerations that are associated with their use, and ways to use ITSs for educational research. Instructor considerations include the time necessary to plan prior to constructing an ITS, the process of constructing ITS lessons for use by students, the method in which students will interact with the ITS, approaches to incorporating ITS use into classes, and the information that instructors would find useful to be output from the ITS. Specifically, the Generalized Intelligent Framework for Tutoring (GIFT), an open-source, domain independent ITS framework will be discussed as an approach to creating adaptive tutoring content for classroom use. GIFT includes straightforward authoring tools for instructors and Subject Matter Experts (SMEs). These authoring tools are powerful, do not require a background in computer science to use, and result in fully adaptive computer-based lessons. Additionally, GIFT provides the flexibility for instructors to bring their pre-generated and already existing instructional material to the system and use it to create ITS lessons. The authoring tools allow the instructor to determine the path of their lesson and the components that their students will experience (i.e. surveys, quizzes, lesson materials, videos). The paper includes details about the development of an instructor dashboard in GIFT, ways for an instructor to use GIFT for educational research, and a discussion of general output information from ITSs that would be relevant to instructors.

Keywords: Intelligent Tutoring Systems, Classroom, Generalized Intelligent Framework for Tutoring, Educational Research, Adaptive Tutoring

1   Introduction

Intelligent Tutoring Systems (ITSs) provide an opportunity for instructors to create adaptive content that their students can engage with as a supplement to their courses. In the current education landscape, even lecture based courses often times have a website associated with them that allows students to download course material, en-

gage in discussions and receive grades. Providing student access to web-based ITSs through these websites is a natural next step. ITSs offer benefits such as personalized adaptive learning, and have been shown to be as effective as a human tutor [1]. One of the major benefits of ITSs, is that unlike a human tutor they can be easily accessed at all hours of the day, and do not get tired or frustrated. ITSs can adapt based on prior knowledge of the individual student, individual differences, or within-tutor performance. This method of instruction provides a tailored, personalized version of lesson materials that can include remediation and clarification of topics. Instructors who author ITSs can make the determination on what type of adaptations they want to occur, and which student individual differences/actions they want to use to determine the adaptations of the system. ITSs can have different benefits and uses based on the type of course that is being taught: online, mixed mode, or lecture [2]. In the case of online courses they may be a vital component of the class that provides important material, whereas, in a lecture course they may serve as an independent supplement to the material that is taught in class. Further, some instructors may want students to engage with ITSs on their own, while others may want them to be used in a computer lab environment with the instructor present for clarification or to assist in classroom management [2].

2   Intelligent Tutoring Systems in the Classroom

Research has shown that ITSs can have positive impacts on learning in a number of different educational domains [1, 3, 4]. Additionally, ITSs can be used either as an added supplement to teaching, or as a component of the classroom. The advantages of ITSs include that they can be used on the student's own time, allow for remediation as needed, and can be engaging as well as motivational. However, the time spent creating materials and remediation for an ITS, is likely to impact the overall adaptivity and outcomes of the ITS. A more adaptive ITS will require more time spent on authoring alternative methods of teaching the required concepts. For instance, if there are 10 different remediation options available to the system based on one concept it will be more adaptive than if there were only 3 pieces of material available. However, authoring this material and considering the situations in which it will be presented does add to the instructor's workload.

While ITSs are a computer-based medium, they can be utilized in both traditional in-person lecture courses, as well as online courses. They can even be beneficial in reduced-seat mixed mode courses. In lecture-based classes, ITSs may be used to provide review and remediation of material that was previously taught, potentially as a review prior to a test. In online classes, ITSs may be one of the primary ways of presenting materials to the students. Mixed-mode classes could potentially integrate ITSs by providing ITS experiences related to the specific material prior to in-class lectures, in order to provide a foundation and context for the material to be learned. ITSs can be useful for not only providing information to students, but also it could be advantageous for students to learn how to create their own ITSs. By planning and creating

ITSs about specific concepts students can reflect upon the material, as well as learn about the functions of these adaptive systems [2]. The utilization of an ITS in these different environments can also provide meaningful output to instructors that can be compiled in the form of a dashboard and be leveraged so that they can adapt their teaching methods as needed. Additionally, ITSs and generalized ITS frameworks can provide a means to perform educational research and examine the impact of different adaptations and interventions within the ITS.

3   Intelligent Tutoring Systems in Educational Research

The use of ITSs as a tool in the classroom has continued to increase throughout the years in U.S. schools. For example, Cognitive Tutor by Carnegie Learning was used in over 2,600 U.S. schools as of 2010 [3]. ITSs have been used for a variety of different age levels spanning from kindergarten to college students. Further, there have been many different ITSs developed in domains as wide-ranging as algebra, physics, medical physiology, law, language learning, and meta-cognitive skills [4]. Comprehensive research examining the effectiveness of ITSs can be found in recent meta-analyses. These meta-analyses examined the effectiveness of ITSs as compared to the effectiveness of typical classroom instruction (i.e., large group and small group human instruction), individual human instruction (i.e., one on one human tutoring), individual computer based instruction (i.e., non-adaptive/intelligent tutoring lacking student/learner modeling), and when the student interacted with an individual textbook [4].

As a tool used in the classroom, ITSs track students' domain knowledge of a subject, learning skills, learning strategies, emotions, or motivation through learner modeling. Further, Steenbergen-Hu and Cooper [3] identified the actions of an ITS as the delivery of learning content to students, tracking and assessing of students' learning progress and adapting to said progress (or lack thereof), and the delivery of appropriate feedback to students. ITSs in the classroom are considered to be superior to traditional computer-based training (CBT) and computer-assisted instruction (CAI) in that ITSs afford unlimited interactions between the ITS and the learner [5].

Steenbergen-Hu and Cooper [3] conducted one of the first meta-analyses examining the effectiveness of math ITSs among K-12 students. The meta-analysis included samples from 1997 to 2010 which had information regarding achievement level, learning outcomes, and an independent comparison group. Overall, their findings suggested that ITS had no negative impact on learning, but only a small positive effect on K-12 mathematical learning was revealed as compared to regular classroom instruction [3]. However, effectiveness of ITSs was greater when compared with homework or human tutoring (i.e., effect sizes of ITS ranged from .20 to .60) [3].

Although small effects were revealed for the effectiveness of ITS on mathematical learning for K-12 grade students, the meta-analysis revealed robust findings to support the use and development of ITSs. Two interesting findings of the meta-analysis

were that shorter uses of the ITS were found to be more effective than long term uses, and that low achievers did not benefit as much from an ITS as other students [3]. These results suggest that individual differences and the length of the exposure to the ITS may have an impact on learning outcomes.

An additional meta-analysis by Ma, Adesope, Nesbit and Liu [4] compared effect sizes from ITS studies that included students of different grade levels, different ITS topic areas, and the way that the ITS was incorporated into the learning environment. In general, ITSs were found to be more effective than standard computer based learning and large lecture classes. The ITSs were effective regardless of how they were incorporated into class (i.e. as a primary means of instruction, as a supplement to material, or an aid). However, ITSs still were not as effective as human one-to one tutoring. These results are insightful, as they show that ITSs may important components of a classroom environment, but the approaches taken with their integration into the classroom should be carefully thought out to ensure that their use is optimized. It was revealed by Ma et al. [4] that the domains of humanities and social sciences are the greatest beneficiaries of ITS use with an effect size of .63. In their meta-analysis, chemistry was the only domain that did not reveal a significant nor moderate effect size.

Although ITSs continue to demonstrate positive achievement outcomes over traditional instruction across a variety of subject domains and education levels, research questions still remain in the use of ITSs and how ITSs can address educational research questions. Also, there are recommendations that ITS researchers can follow in reporting and documenting their results to improve the overall ability for ITS researchers to draw more consistent and reliable conclusions from reported research. Steenbergen-Hu and Cooper [3] found ITSs to have a greater impact on moderate achieving students than low achievers. There is a need to examine how ITSs can better impact the learning outcomes of the students that need it the most. How can ITSs be leveraged to affect students of different and lower achieving levels? Further, research using ITSs should examine and develop a better understanding of why higher achieving students benefit more from the use of ITSs. It is not unlikely to hypothesize that lower achieving students may have less motivation than higher achieving students. Therefore, how ITSs better leverage intrinsic and extrinsic motivational factors is an example of a research question worth further pursuing.

As pointed out by Ma et al. [4], although ITSs have demonstrated effectiveness, it is still difficult to definitively come to a consensus on explanations for the effectiveness of ITSs. Further research is necessary to address and offer explanations for why ITSs are effective in order to improve the development of ITSs. Also, this research should provide further insight on how to improve the efficacy of instructors in the classroom.

Lastly, there are some recommendations researchers can adhere to when reporting and documenting the results of their research in order for others to draw more reliable and consistent conclusions from the reported research. Researchers should adhere to the standards of reporting basic statistics such as means and standard deviations, and Ma

et al. [4] recommend development of a taxonomy of ITS design. Developing a taxonomy of ITS design would enhance the standardization of ITS research reporting, ideally resulting in quicker ITS research advancements and a common framework for researchers and practitioners to draw reliable and valid conclusions for the use of ITSs in education.

4 The Generalized Intelligent Framework for Tutoring (GIFT) and Educational Research

In order to study the effectiveness of ITSs, an ITS not only has to be created, but researchers must put together carefully constructed experiments to determine the real world application and benefits of ITS use. Different approaches can be taken to conducting educational research with ITSs. Comparisons can be made between grades from students who were in a previous ITS-less versions of the course as opposed to an ITS-enhanced version. Pre and post test can be given before and after ITS use. In an online class, the pre and post performance of students that engaged with an ITS can be compared to those who just received non-adaptive computer-based material. One of the inherent difficulties with designing a study that actively uses students and provides different means of providing material to them is that the instructor does not want to offer more of an advantage to one student over another by providing them with better instructional materials. Therefore, it is important to carefully design the materials to make sure that they are equivalent in content. The time the student spends with the material can also be a metric to examine, as those with the ITS may more efficiently peruse the material as opposed to receiving a regular all inclusive version.

While meta-analyses were able to compare overall effect sizes for ITSs, they do not allow for direct comparisons between ITSs of different subject types in controlled experimental fashion. If ITSs in different subject areas were constructed using the same framework and with consistency, then perhaps their learning outcomes can be more directly compared to each other. For instance, are there more learning gains when ITSs are used for algebra as opposed to when they are used for learning a language? Further, an area that has not received as much attention is the components of the learner model that are tracked during interaction with the ITS or that result in adaptations [6]. Research could further investigate these questions by engaging in experiments that vary the individual differences or characteristics that adaptation occurs based on. For instance, is there an improved outcome to adapting based on prior knowledge and motivation level in an algebra tutor, or is it more advantageous to adapt just based on prior knowledge? Generalized frameworks for ITSs can help offer an opportunity to research these types of questions.

Most ITSs are tightly coupled with the material that they are teaching, and are not reusable. However, the Generalized Intelligent Framework for Tutoring (GIFT), is a domain-independent ITS framework. The tools that exist within GIFT can be used to create adaptive tutoring in any subject or topic. Due to this, it allows for reusability of

material and adaptability of the ITS without needing to start from scratch or develop an entirely new framework. GIFT is made up of different modules and components: the learner module, pedagogical module, domain module, sensor module, gateway module, and the tutor-user interface [7]. The only module that is tied to the domain content is the domain module. The flexibility that exists within GIFT also allows for changes to be made to the types of information that is being tracked in the learner module, the types of adaptations that are recommended by the pedagogical module, as well as the material in the domain module. Therefore, GIFT provides an opportunity to examine the impact of changing the selected characteristics and representations within these modules without needing to dramatically change or reprogram an already established ITS. This functionality opens up opportunities for further expanding educational research using ITSs. These types of research questions allow for educators to research what the optimal individual differences to adapt to are, as well as if there are advantages to one type of adaptation over another type. The information gathered from this can then be applied in educational environments whether they are online or in in-person classes. Additionally, the flexibility of GIFT allows for instructors to utilize the elements of it in the classroom to add to and enhance the way that they interact with their students.

5   Applying GIFT in the Classroom

While much of the research conducted to enable GIFT as an adaptive training tool has been focused on standalone (no human-in-the-loop) one-to-one tutoring capabilities, the goal has always been to have GIFT used in a classroom environment as an aid to human instructors too. This section discusses the information needs of human instructors which would enable them to evaluate and manage concurrent computer-based tutoring sessions of their students. We examine what information the instructor might need to optimize decisions about when and where they allocate their time to intercede with students who need help beyond what a computer-based tutor is able to provide. We begin by discussing what information about the student is already available to GIFT-based tutors and later extend this model to support the classroom paradigm.

The learner model in GIFT-based tutors includes information from various sources. As noted in the various updates of the Learning Effect Model (LEM) [8, 9, 10], this information originates from five primary sources: 1) real-time student interaction with the tutor and the training environment (e.g., responses to requests for information); 2) real-time sensor data and physiological states based on sensor data; 3) historical data from record stores which include demographics, domain experiences, knowledge, achievements and the results of validated assessments (e.g., grit surveys, personality and other trait appraisal instruments); 4) real-time assessment of performance based on learner progress toward learning goals and other behavioral states based on sensor

data; and 5) external environments (e.g., entity level data from a simulation integrated with a GIFT-based tutor through a standardized GIFT gateway).

GIFT uses this information to select strategies and implement instructional tactics with the goal of accelerating and optimizing learning, performance, retention, and the transfer of skills developed during training to the work or operational environment. A consideration in developing a dashboard (information resource) for application in a classroom is the migration of each student from one quadrant (i.e., rules, examples, recall, and practice) to the next as described by Merrill's [11] Component Display Theory and implemented in GIFT. In the classroom use case, GIFT should be able to provide a comprehensive picture of the student population at a glance so the instructor can decide where to allocate their time in support of student learning objectives. This could mean alerting the instructor when students struggle with domain concepts and content or when they fall below expectations based on past performance.

Bull and Nghiem [12] and Guerra, Hosseini, Somyurek and Brusilovsky [13] recommend an open learner modeling approach which is designed to help learners to better understand their learning processes with a model which is accessible to the student, the instructor, and their peers. Bull and Ngheim [12] also note the following benefits of the open learner modeling approach: 1) improves the accuracy of the learner model by allowing students to contribute information to it; 2) promotes reflection; and 3) helps the tutor plan and monitor learning based upon the foundation of information available in the learner model. The information available in an open learner model ranges from performance statistics (e.g., quiz grades) to progress toward goals (e.g., completed 58% of assigned work). Guerra et al [13] suggest a graphic visualization of the learner's activities (e.g., quizzes, examples) and domain topics in their mastery grids system which uses various shades of green to indicate student performance, shades of blue to represent reference group performance, and a combination of green and blue to indicate how an individual student is performing with respect to the reference group. This system allows a student, instructor or peers to quickly assess their performance in a variety of activities and topical areas.

Considering the open learner model and various states and traits available within the GIFT architecture, we recommend a hybrid system to allow instructors to address not only performance concerns, but also the affective state, domain competency, and learning readiness of their students. A simple dashboard (Figure 1) might show a classroom of 20 student icons color coded to show the instructor the overall state of the student. Students with green status (e.g., Students A, C and E) are on track in the pursuit of their learning goals and are not currently experiencing any negative affective states. Students with yellow status (e.g., Students D, F and H) may be performing slightly below expectation based on their domain competency and/or experiencing negative affective states relative to learning readiness. Students with red status (e.g.,

Student B) may be significantly underperforming or experiencing negative affective states that significantly curtail learning. Finally, white squares represent neutral status which may mean that the student has not yet begun the set of tasks in the domain under instruction.
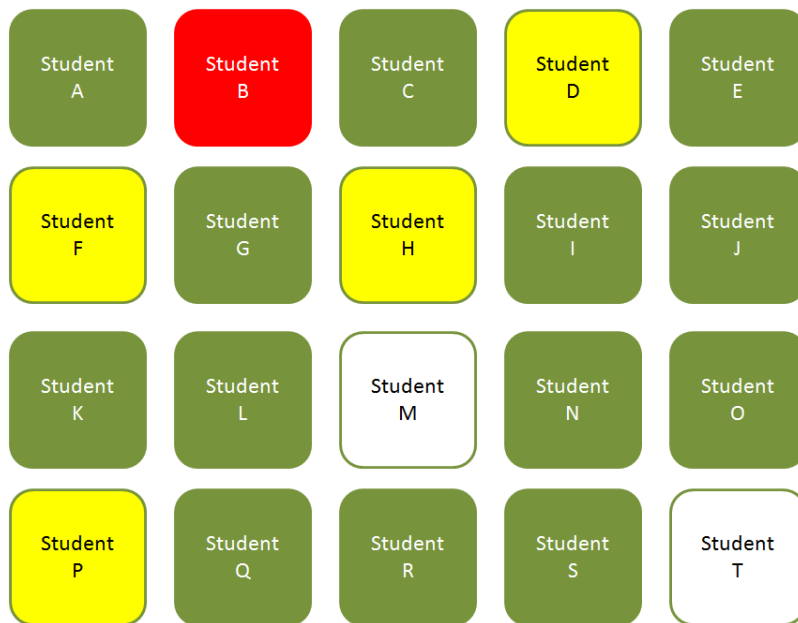

Fig. 1. Top level view of notional GIFT Dashboard

Details about any of the students represented in this dashboard may be viewed by clicking on the appropriate student icon. Figure 2 shows an example of the status of an individual student. There is a breakdown of status based on concept, affective state, and quadrant based on Component Display Theory. The same color scheme as the top level dashboard view is used.
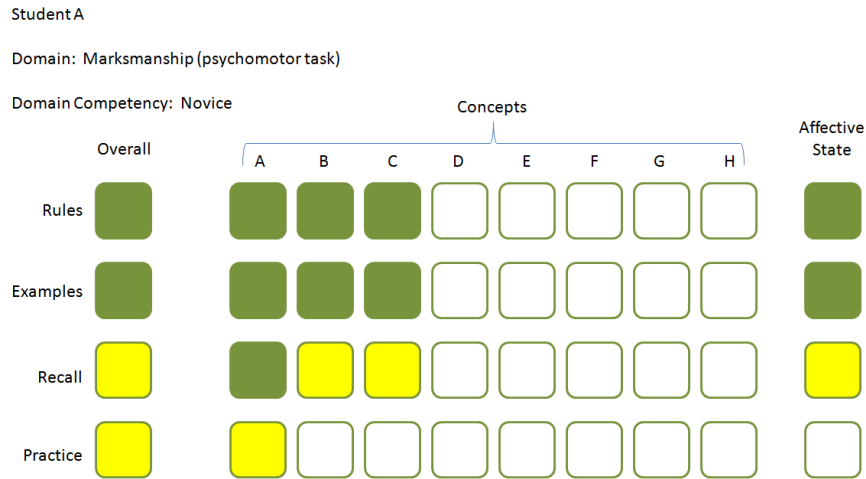
Fig. 2. Student detail level view of notional GIFT Dashboard

6 Future Considerations and Recommendations for GIFT to assist in Educational and Classroom Use

ITSs may seem superficially similar to linear, computer-based training (CBT). However, ITSs adapt to the profile of a learner, which can include their current and prior experiences and performance, learning preferences, affective states, and so on. Thus the resources, authoring, and pre-production required in order to build an effective tutor are greater than that of computer based training. GIFT, as an intelligent tutoring platform, intends to provide the means to create, deploy, and manage adaptive training content while lowering the skill and resource barriers to accomplishing those tasks. While great progress has been made in service of those core principles, there remain opportunities for improvement. Here, we will describe considerations and recommendations for future research, design, and development in GIFT supporting classroom education and educational research, along the dimensions of *authoring, instructional support, and research management*.

6.1 Considerations and Recommendations for Classroom Education: Authoring

The concept of creating a tutor is a relatively new content creation paradigm. Therefore, one of the greatest challenges to tutor authoring is how to best cultivate mental models of ITSs in novice end-users, and cultivating an authoring user experience for users that encourages the creation of truly adaptive tutors (as opposed to producing linear CBT). GIFT currently provides a series of authoring tools, intended to reduce the time and skill required to produce tutors. Our current approach in developing a user experience for tutor authoring is based upon tenants of mental model theory:

when confronted with a new system, individuals will rely upon mental models of systems perceived to be familiar to the new system [14]; and that mental models help make sense of the form, function, and purpose of a system [15].

With that in mind, GIFT's current authoring tools use interfaces and interaction paradigms that are intended to look and feel familiar to other productivity tasks such as building a flow chart, filling out a form, or creating a web-page. The idea is that familiar interface elements from other productivity applications will help to form the foundation of a mental model for tutor authoring. Much of this effort has been targeted at the core elements of the authoring experience (e.g., sequencing elements, adding media, developing survey material) as well as quality-of-life improvements (e.g., auto-save, copy/paste, minimizing clicks and pop-ups) [16].

With a system that is reasonably learnable and usable, we are discovering new considerations for education with an expanded user base. Particularly, many authors bring their existing content to GIFT (or any ITS), however this content is largely not in a format suitable for adaptation. That content is generally intended to be viewed in its entirety by all of the learners, constituting CBT. While GIFT is not a media creation tool, future GIFT development should support the semi-automated process of content generation and/or formatting for adaptation based on learner characteristics. For example, that might involve assisting the author in sub-diving an existing slide show or print material into core, remedial, and advanced content and then placing that content in the appropriate course elements within a GIFT course. Or, authoring support may take the form of intelligently interfacing with external content repositories to help locate and suggest additional content to the author to include in their tutor.

Future GIFT-related research should consider novel ways to provide adaptations beyond content selection. GIFT, for instance, presents tutors within its own custom tutor-user interface (TUI). Improvements to the TUI could be made, configurable via the authoring tools, which would provide certain overlays and interface elements that would change and/or appear based on the learner's profile. For example, a learner that is highly competitive may be presented with the option to view a leaderboard in an effort to build motivation, but such a TUI element would not be shown if the system believes it would only demoralize that learner. The actual learning content remains unchanged. Leaderboards, specifically, come from a larger class of TUI elements inspired by gamification [17], however, there are other ways in which existing media content can be enhanced or modified through the TUI, such as options for background music, context personalization [18], or the ability to customize the tutor avatar with which a learner interacts.

6.2 Considerations and Recommendations for Classroom Education: Instructional Support

As described in Section 5, ITSs have the potential to produce large amounts of data, including those about the learner (e.g., profile, sensors, preferences), the learner's

interaction within the ITS (and linked, external practice environments), actions taken by the ITS based on the learner model, as well as the learning content and assessments presented to each individual learner. Data sources may also include information external to GIFT, such as a learner record store [19]. With respect to instructional support, the primary consideration for GIFT is to provide a *dashboard* that enables instructors to quickly perform data exploration and high level analyses in order to ascertain the health of the class, and make decisions regarding interventions for high or low performing students.

Given the nature of a flexible, adaptive system like GIFT, there may not be a single best solution for a dashboard. Each row of student data within the same course may contain different columns of information, depending on the adaptive paths encountered within the tutor. Since GIFT is a domain-independent platform, the types of data that are generated across courses will vary wildly. Further, different instructors in different courses may need to answer different types of questions regarding their courses, suggesting that there may not be a single user experience that best fits all these cases. To that end, GIFT should consider the perspective that adaptive tutoring systems will require adaptive instructional dashboards.

The high-level notional concepts presented in Figures 1 and 2 (above) help to answer questions regarding *how* the students in the class are performing, and those views may remain fairly consistent across GIFT modules. As an instructor drills down into the data however, customizable views will be required to help answer questions about *why* the students are exhibiting certain levels of performance [20]. Again, the data available to answer these questions depends upon the unique composition of the tutor. Therefore, a user-centered design strategy should be followed in pursuit of a GIFT instructor dashboard. Operationally, a modular dashboard should be built around instructors' work goals, and the associated tasks required in order to meet those goals (Figure 3). Specifically, GIFT would provide semi-automated support to the instructor in constructing figures and charts, and the instructor should be able to organize those reports into a customizable view, similar to the interface of an analytics dashboard for website usage.

Consider a use case illustrated in Figure 3. Using the dashboard, an instructor notes that one student is performing poorly in a course, relative to the performance of the other students. Note that Figure 1 is one of the views that the instructor has added to their custom dashboard. On the surface, the student appears to be engaging with the tutor, and the course materials contained within, but the instructor wants to investigate the low-performing student's actions within the system in greater detail. Using a modular instructor dashboard in GIFT, the instructor decides to begin examining the extent to which students of different performance levels interact with various types of instructional media contained within that lesson. From a list automatically-populated of available charts, figures, and tables, the instructor adds the relevant module to their dashboard view, and selects three students for comparison. The instructor notes that the low-performing student appears to have spent the same amount of time viewing

the lesson material with the exception of some of the image content. The instructor can now investigate whether the low performing student missed important information contained within some of the images in the lesson. Data exploration can continue in this way to corroborate this potential linkage between the student's performance and the time spent with a particular type of lesson material.
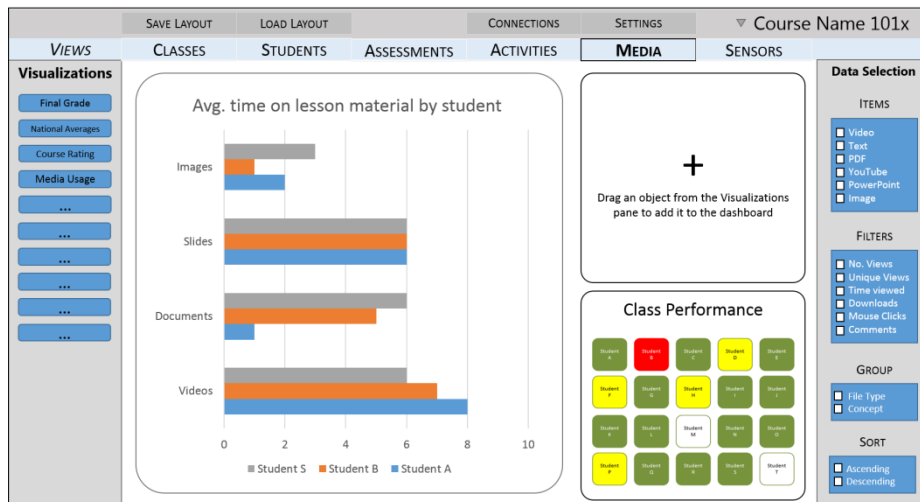

Fig. 3. Conceptual mock-up of a modular, semi-automated instructor dashboard for GIFT

Functional considerations should also be made to improve the usability of the dashboard tools. Layouts and configured visualizations should be able to be saved as views, for use in future courses, or to share with other instructors. Dashboard elements should be interactive: Hovering the cursor over individual data points should provide pop-ups with additional details. Clicking on a relevant data point, such as "Student A" in the Class Performance visualization in Figure 3, should produce the view found in Figure 2, by "zooming into" that view as an underlying element. Elements should be movable, resizable, and support common productivity functions such as cut, copy, and paste. Similar to the authoring tools UX, overall quality-of-life improvements will help to make the tools more efficient and allow the instructor to spend less time setting up the dashboard, and more time exploring the data [21].

6.3 Considerations and Recommendations for Education Research: Management

GIFT has been used for research purposes since its inception, and it is upon research that GIFT's pedagogical engine and other features are based [22] GIFT has only recently, however, been updated with features directly supporting tasks associated with preparing, administering, and managing research. Currently, core functionality is in place that allows an existing GIFT module to be spawned into a "research version" of

that module [16]. Doing that creates a non-editable version of the module, with the intent of maintaining the consistency of the trials across participants. A unique URL is generated that allows participants to directly access the course without a GIFT Account, with the intent of protecting the anonymity of their data. Access to the study can be paused and resumed in accordance with data collection timelines and regulatory bodies. GIFT's research tools also provide interfaces for downloading customized data files and reports when desired.

Future considerations for GIFT in support of educational research could include explicit features for creating and managing treatments/manipulations within *experimental* versions of the material to be learned, as well as the distribution of participants into those sets of materials. Consider a use case in which a researcher wants to implement three versions of a educational material covering a concept that only differ by a specific element. The researcher also wants to semi-randomly distribute participants into the three versions of the material, but ensure that each cell has equal participants with similar distributions of high/low motivated learners. GIFT might handle this use case in one of two ways, either internally or externally to the course. One implementation would use the same overall GIFT course with a special course element containing all three versions and logic for specifying the distribution into the permutations. Alternatively, three separate versions of the material could be somehow "linked" together in a way that version control is maintained across them with the exception of the elements intended to be manipulated. Randomization and assignment of participants would then be handled through the top-level Research UI of the GIFT interface. Determining the "best" design implementation for this functionality may come down to preference, as the design of adaptive tutors themselves is still evolving.

Finally, more robust reporting tools are needed for educational research using adaptive tutors. GIFT is intended to be a flexible, domain independent platform, therefore the types of tutors that can be created will vary wildly. GIFT also adapts on a number of learner characteristics using both discrete-time, outer-loop logic as well as real-time (or near real-time), inner-loop logic. Sources of learner data may also come from various sources (described earlier in this work). It logically follows that the data outputs from educational research will require different reporting formats as well beyond the current capabilities of the reporting tool currently provided by the GIFT web-application. Instructor dashboards, described in the prior section, may assist the researcher as well in conducting exploratory analyses with partial or complete data sets.

7 Conclusions and Recommendations

ITSs can be extremely useful to instructors of courses, regardless of the modality. They have the ability to engage students with material that may have been missed, or that was not completely understood. Additionally, they are adaptive to the individual such that the prior knowledge and performance of the student will impact the material

that they are provided. ITSs have been demonstrated to be useful in both the laboratory environment as well as in classroom environments [23, 24, 25].

There are many educational research questions that can still be examined in ITSs, such as what the ideal components of the learner model are, a comparison in effectiveness of ITSs between domains, and the impact ITSs have when implemented in an in-person vs. an online course. A domain-independent ITS framework such as GIFT provides opportunities to construct ITSs to contribute to the answers to these questions, and to enhance the classroom experience. It is recommended that GIFT be used to pursue these and similar research questions that are not practical or able to be asked in traditional ITSs. As GIFT and other ITSs continue to be developed for both practical use and educational research, it is recommended that instructor dashboards are designed to be customizable and provide a way to harness the rich data that is available from ITSs about student performance, states, and progress. ITSs can be extremely useful to instructors, and can be incorporated into classes in a number of different meaningful ways including as a means to: provide information, remediate information, monitor student performance/state, and to conduct educational research.

8 References

1. VanLehn, K.: The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. Educational Psychologist, **46(4)**, 197-221 (2011).
2. Sinatra, A.M., Sottilare, R., Ososky, S.: Assessment in Intelligent Tutoring Systems in Traditional, Mixed Mode, and Online Courses, In: Design Recommendations for Intelligent Tutoring Systems, Volume 5: Assessment (in press).
3. Steenbergen-Hu, S., Cooper, H.: A meta-analysis of the effectiveness of intelligent tutoring systems on K-12 students' mathematical learning. Journal of Educational Psychology, **105(4)**. 970-987 (2013).
4. Ma, W., Adesope, O. O., Nesbit, J. C., Liu, Q.: Intelligent tutoring systems and learning outcomes: A meta-analysis. Journal of Educational Psychology, **106(4)**, 901-918 (2014).
5. Graesser, A. C., Conley, M., Olney, A.: Intelligent tutoring systems. In K. R. Harris, S. Graham, & T. Urban (Eds.), APA educational psychology handbook: Vol. 3. Applications to learning and teaching, 451–473. Washington, DC: American Psychological Association (2011).
6. Holden, H. K., Sinatra, A. M.: The Need for Empirical Evaluation of Learner Model Elements. In Design Recommendations for Intelligent Tutoring Systems, Volume 1: Learner Modeling, 87 (2013).
7. Sottilare, R. A., Goldberg, B. S., Brawner, K. W., Holden, H. K.: A modular framework to support the authoring and assessment of adaptive computer-based tutoring systems (CBTS). In Proceedings of the Interservice/Industry Training, Simulation, and Education Conference, 1-13 (2012).

8. Sottilare, R.: Considerations in the development of an ontology for a Generalized Intelligent Framework for Tutoring. International Defense & Homeland Security Simulation Workshop in Proceedings of the I3M Conference. Vienna, Austria (September 2012).

9. Sottilare, R., Ragusa, C., Hoffman, M., Goldberg, B.: Characterizing an adaptive tutoring learning effect chain for individual and team tutoring. In Proceedings of the Interservice/Industry Training Simulation & Education Conference, Orlando, Florida (December 2013).

10. Sottilare, R.: Elements of a Learning Effect Model to Support an Adaptive Instructional Framework. In R. Sottilare & S. Ososky (Eds.) 4th Annual GIFT Users Symposium (GIFTSym4) Army Research Laboratory, Orlando, Florida (2016) ISBN: 978-0-9977257-0-4.

11. Merrill, M. D.: Component display theory. Instructional-design theories and models: An overview of their current status, **1**, 282-333 (1983).

12. Bull, S., Nghiem, T.: Helping learners to understand themselves with a learner model open to students, peers and instructors. In Proceedings of workshop on individual and group modelling methods that help learners understand themselves, International Conference on Intelligent Tutoring Systems, 5-13 (2002).

13. Guerra, J., Hosseini, R., Somyurek, S., Brusilovsky, P.: An intelligent interface for learning content: Combining an open learner model and social comparison to support self-regulated learning and engagement. In Proceedings of the 21st International Conference on Intelligent User Interfaces, 152-163. ACM. (March 2016).

14. Craik, K.: *The nature of explanation*. Cambridge: Cambridge University Press. (1943).

15. Rouse, W. B., Morris, N. M.: On looking into the black box: Prospects and limits in the search for mental models. Psychological Bulletin, **100(3)**, 349-363 (1986). doi: 10.1037/0033-2909.100.3.349

16. Ososky, S.: Designing the user experience of the GIFT Cloud authoring tools. In R. Sottilare & S. Ososky (Eds.), In Proceedings of the 4th Annual GIFT Users Symposium Volume 4, pp. 145-156. Orlando, FL: U.S. Army Research Laboratory (2016).

17. Ososky, S.: 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, 640-651: Springer International Publishing (2015).

18. Sinatra, A. M.: A Review of Self-Reference and Context Personalization in Different Computer-Based Educational Domains. Design Recommendations for Intelligent Tutoring Systems: Volume 4-Domain Modeling, 107 (2016).

19. Sabin, M. (2012). Student-pull instead of instructor-push: in preparation for a student learning dashboard. Journal of Computing Sciences in Colleges, **27(6)**, 70-72 (2012).

20. Siemens, G., Gasevic, D., Haythornthwaite, C., Dawson, S., Shum, S. B., Ferguson, R., . . . Baker, R. Open Learning Analytics: an integrated & modularized platform. Open University Press Doctoral dissertation (2011).

21. Lightbown, D.: Designing the User Experience of Game Development Tools. Boca Raton, FL: CRC Press (2015).

22. Wang-Costello, J., Goldberg, B., Tarr, R. W., Cintron, L. M., Jiang, H.: Creating an advanced pedagogical model to improve intelligent tutoring technologies. In The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC) (2013).

23. Corbett, A. T., Koedinger, K. R., Hadley, W. H.: Cognitive Tutors: From the research classroom to all classrooms. Technology enhanced learning: Opportunities for change, 235-263 (2001).

24. Koedinger, K. R., Aleven, V., Roll, I., Baker, R.: In vivo experiments on whether supporting metacognition in intelligent tutoring systems yields robust learning. Handbook of metacognition in education, 897-964 (2009).

25. Yacef, K.: The Logic-ITA in the classroom: a medium scale experiment. International Journal of Artificial Intelligence in Education, **15(1)**, 41-62 (2005).