Tech., Inst., Cognition and Learning, Vol. 11, pp. 5–38
Reprints available directly from the publisher
Photocopying permitted by license only

A Comprehensive Review of Design Goals and Emerging Solutions for Adaptive Instructional Systems

ROBERT A. SOTTILARE

US Army Research Laboratory, Human Research & Engineering Directorate,
Associate Director for Training Technology (RDRL-HRR), USA,
e-mail: robert.a.sottilare.civ@mail.mil

This article is intended as a companion document to the more focused report provided by the author at the 2017 American Education Research Association (AERA) Conference as part of the Technology, Instruction, Cognition & Learning Special Interest Group's Symposium on Intelligent Tutoring Systems (ITSs). Both the AERA talk and this article focus on adaptive instructional systems (AISs) which are comprised of learners, Intelligent Tutoring Systems (ITSs), and external (non-adaptive) instructional environments. AISs tailor instructional experiences for individual learners and teams of learners based on a model of their learning needs and preferences. An exemplar of an AIS is the Generalized Intelligent Framework for Tutoring (GIFT), an open source architecture for authoring, delivering, managing, and evaluating AIS technologies (tools and methods). This article reviews desired states for AISs in the context of enhancements to GIFT capabilities. This article covers a wide range of desired states for AISs and their affiliated design goals, challenges, and emerging solutions. While we consider the review presented in this paper comprehensive, we acknowledge that it is far from exhaustive. Our primary goal is to present the state of art, potential, and practice in ITS design in order to engage the education and training community in our quest to make AISs ubiquitous.

Keywords: adaptive instruction, Intelligent Tutoring Systems, adaptive instructional systems, Generalized Intelligent Framework for Tutoring (GIFT), authoring tools, accelerated learning, learner modeling, domain modeling, automated instructional management, distributed learning, mobile learning

*Corresponding author: robert.a.sottilare.civ@mail.mil

1 INTRODUCTION

Many Intelligent Tutoring Systems (ITSs) are highly effective learning tools and provide individually-tailored instruction in well-defined, cognitive task domains (e.g., mathematics, physics, or software programming). The adaptive instruction provided by ITSs intelligently selects and sequences content, adapts feedback, and guides the learner with the goal of optimizing their learning and performance. In his meta-analysis on the effectiveness of ITSs, VanLehn (2011) notes that ITSs have evolved to parity with expert human tutors. Unfortunately, this parity exists in a very limited set of training and educational domains, usually mathematics, physics or software programming. If we want future ITSs to be ubiquitous instructional tools, there are significantbarriers to overcome.

This article focuses on significant goals, challenges, and emerging solutions related to the development and adoption of ITSs. To this end, we have adopted a model of ITSs as part of a functional toolset for authoring, instructional management, and evaluation that we call an *adaptive instructional system (AIS)*. The AIS concept extends the model of advanced situated tutors proposed by Schatz, Bowers, & Nicholson (2009) by including both runtime (instruction and evaluation) and offline functions (e.g., authoring) in an instructional architecture known as the Generalized Intelligent Framework for Tutoring (GIFT; Sottilare, Brawner, Goldberg, and Holden, 2012; Sottilare, Brawner, Sinatra, and Johnston, 2017). GIFT is being developed with the AIS model in mind and we refer to current and emerging capabilities as desired endstates within GIFT. Within the runtime AIS, we consider models of the learner(s), the ITS as a tool to manage/guide adaptive instruction, and an optional external environment to encourage the reuse of non-adaptive instructional systems (e.g., simulations, games, web pages) as content providers for the AIS.

Functionally, the AIS may be distinguished from traditional ITSs by its interactions. For ITSs, the tutoring agent traditionally interacts with a learner and an internal expert model to assess the learner's progress towards one or more learning objectives. In an AIS, this interaction may also include some external environment as noted above and shown in Figure 1. AISs also are "self-evaluating" in that they review their own performance with respect to their decisions, actions, and policies. Figure 1 illustrates these interactions with both the learner and the environment where the tutor element of the AIS observes both the learner and the environment, applies appropriate policies (e.g., constraints, rules, best practices), and then evaluates the effectiveness of those policies. The observations lead to decisions by the tutor which lead to actions by the tutor, which lead to an evaluation of effectivess of the policy selected and applied, and finally, leads to changes to the policy as required. This is a reinforcement learning system.

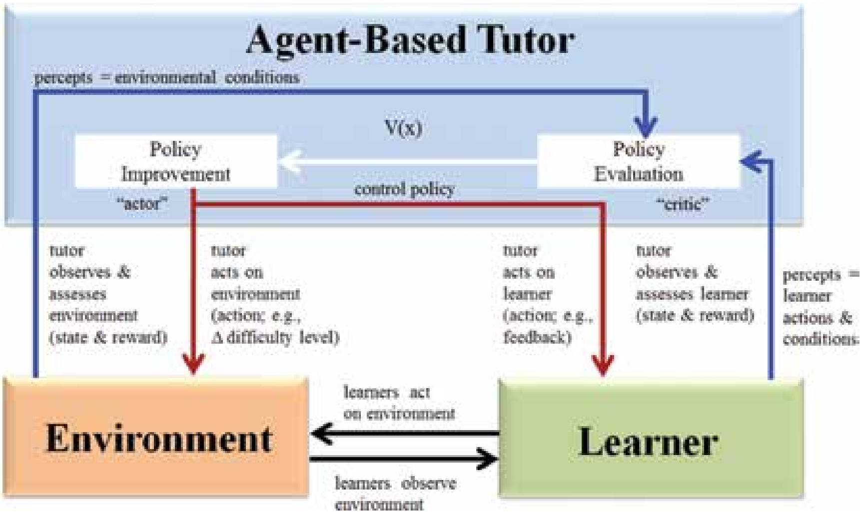

FIGURE 1
Runtime Adaptive Instructional System.

The integration of previously non-adaptive external environments with ITSs provide adaptive instruction to learners based on their individual learning needs and preferences. The runtime AIS can also be configured for experimentation to support the evaluation of learner/instructional/domain models and hypotheses testing through the GIFT testbed functionality as derived from Hanks, Pollack, and Cohen (1993) and described in Figure 2. Learner attributes, domain attributes, and instructional policies, strategies, and tactics can be manipulated to conduct validation experiments to identify the effect of each on individual learning, performance, retention, and transfer for skills from instructional environments to work environments.

In the offline AIS, GIFT authoring processes provide ITS developers with tools and methods to organize course/lesson material, sequence blocks of instruction, create methods of assessment to determine the learner(s) progress toward learning objectives, and associate optimal tutor interactions (e.g., feedback or remediation) identified through experimentation with learner behaviors and progress.

AISs, as a concept, may represent a leap forward in the evolution of ITS design, but realizing the full potential of the concept is frought with challenges. Whereas ITSs are generally standalone systems focused on a single instructional domain, AISs are able to link with non-adaptive external environments (e.g., simulators, simulations, games, Massive Open Online Courses (MOOCs) or webpages) to drive adaptations (e.g., changes to difficulty or learner support level) in

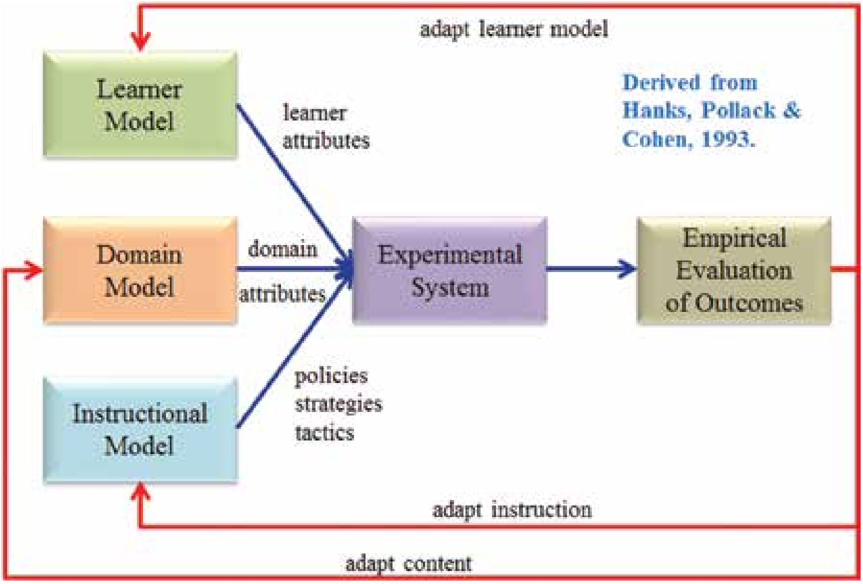

FIGURE 2
Runtime AIS Experimentation in GIFT.

those environments, and they may do so for a diverse scope of instructional domains. With their extended functionality and interaction with external environments, AISs have the potential to be even more complex than the ITSs that are prevalent today. The question is: can AISs be designed with all the benefits of adaptive instruction, but with much less of the costs of traditional ITSs? Based on their complexity, their lack of reusable components, and the expert skillset required to design, create and update them, ITSs are expensive to author. Can AIS provide the tools needed to reduce authoring and integration costs while also lowering the barrier of skills required to create more complex AISs?

A design goal for AISs is to extend adaptive instruction beyond traditional training and educational domains to make AISs more versatile. Today, ITSs are primarily focused on procedural, well-defined domains like mathematics, physics, and software programming, and they are just emerging as instructional tools for collaborative problem solving and team tutoring. The capability to support both individual and group instruction will greatly expand the number of application domains for AISs and make them more valuable instructional tools.

The high degree of effectiveness of ITSs easily make them cost effective in high density domains (e.g., high school mathematics) with large populations of learners. This is exciting news, but several barriers to ITS adoption in new instructional domains remain. It is impractical at this time to develop ITSs for low density instructional domains (e.g., specialized fields with small populations of learners). However, the application of ITSs to a broader set of high density domains should be a goal and affordability is a primary challenge to their adoption. Simpler and more automated authoring processes hold the key to making AISs affordable across a broad set of instructional domains.

What if ITSs were easier to author in a broader range of task domains? What if the return on investment (ROI) made the authoring of even low density adaptive instructional domains cost effective? What does the AIS need to know about the learner(s), domain, and context to make effective instructional decisions? What are the goals, challenges, and emerging solutions related to the development, delivery, and evaluation of adaptive instruction?

To this end we have identified eight goals for AIS enhancement along with their associated challenges and discussion of emerging solutions: 1) developing efficient authoring processes, 2) developing effective instructional decisions, 3) modeling learner trends 4) building rapport and engagement with learners, 5) modeling collective instructional domains, 6) expanding adaptive instruction to a broader array of task domains, 7) evaluating the effectiveness and efficiency of AISs, and 8) supporting distributed/mobile learning. Woolf (2010) identified seven grand challenges for education technology. We discuss four of these challenges: personalizing learning, assessing learning, supporting social learning, and diminishing barriers, and others in the context of their relationship to our eight AIS enhancement goals.

2 GOAL #1: DEVELOPING EFFICIENT AUTHORING PROCESSES

Authoring is the process of gathering, organizing, and sequencing content for delivery to the learner. Part of the authoring process is also identifying learning objectives (known as *concepts* in GIFT) and associating content, learner attributes (states and traits), and measures of learning and performance with those learning objectives to allow AISs to track learner progress. AISs and their major adaptive element, ITSs, are expensive to develop because it takes a set of very specific skills and a keen understanding of intricate instructional processes to build them. ITSs are often purpose-built (domain-specific) systems built by teams whose expertise usually includes instructional design, software programming, human factors, and extensive domain knowledge (e.g., subject matter experts).

One of the primary challenges to making AISs practical for use by the masses is improving their efficiency by reducing the skill and time required to author/

create them. Woolf (2010) discusses the challenge of diminishing barriers. In the authoring process, this can be interpreted as enabling authors with varying knowledge and skills to be successful at the authoring task. Toward this end, a set of associated authoring goals were developed by Sottilare, Goldberg, Brawner, & Holden (2012) for GIFT as adapted from Murray (1999, 2003). Most of these goals target efficiency in the authoring process:

- Decrease the resources (materials, time, cost, etc.) required to author an ITS
- Decrease the skill threshold required by various user groups associated with authoring and managing an ITS as part of a curriculum
- Enable rapid prototyping of intelligent tutors for rapid design and evaluation of capabilities
- Develop/adopt standards, including common tools and interfaces, for tutor authoring
- Promote reuse of content, modules, and data structures in tutors

An objective of adaptive instruction is for each learner to have customized/tailored learning experiences based on their prior domain knowledge, goal orientation, and other personalization factors to engage them in each and every learning experience. To this end, an AIS must have multiple types of content/scenarios to present to a variety of learners at runtime. Murray (2003) estimated that the authoring of non-adaptive computer-based instruction requires 100–300 hours for a team of skilled computer programmers, instructional designers, and subject matter experts time to create 1 hour of non-adaptive computer-based instruction. AISs require more content (e.g., presentations, media, question banks, conversation trees, simulation scenarios, assessments, and instructional strategies) to provide a variety of adaptive paths based on individual differences and this increases the effort to author and its associated tasks of developing/finding and organizing content. Two approaches are being pursued concurrently to make AISs easier to author: 1) improving the usability of authoring tools to make the authoring process less complex for authors, and 2) automating parts of the authoring process to reduce/eliminate the author's workload. 2.1 Improving the Usability of GIFT Authoring Tools.

Specifically, in GIFT, we are attempting to improve the authoring tool interfaces to make the process more natural for the author. This has made the GIFT authoring process more efficient by reducing the author's cognitive load. Cognitive load was reduced by eliminating extraneous information and only presenting controls and information that are pertinent to the authoring task (germane) or pertinent to learning the task (instrinsic). We are also attempting to reduce/eliminate steps in the authoring process by providing tools or job aids.

2.1.1 Reducing Cognitive Load through Simpler Authoring Interfaces

Since AISs consider the learner to be an integral part of the system upfront, usability is always a consideration and one of the primary learners in an AIS is the author. The author comes to the task of creating an instructional system with a set of skills that may not include software programming or instruction design, but usually comes with some knowledge of the domain. GIFT attempts to overcome the author's deficiencies by eliminating the need programming to a large degree and baking the principles of instruction into the process to guide the learner in developing effective instruction. While programming is required to join new external systems (e.g., training environment or sensor) to GIFT, once a gateway is created for an application, the application can be used by dragging and dropping a representative object into the learning sequence for any GIFT course (Figure 3).

The drag-and-drop interface allows authors to select and sequence course objects (e.g., various types of media content, external applications, conversation trees, adaptive courseflow objects, and surveys). Once the author selects and sequence the object, GIFT provides a window for the author to configure the object with the information needed for it to be functional during instruction. This provides the author with a set of templated objects with which they can build courses.

GIFT integrates instructional design principles primarily through the adaptive courseflow object which incorporates Merrill's Component Display Theory (CDT) of instruction (1983). This courseflow object sequence and loops the

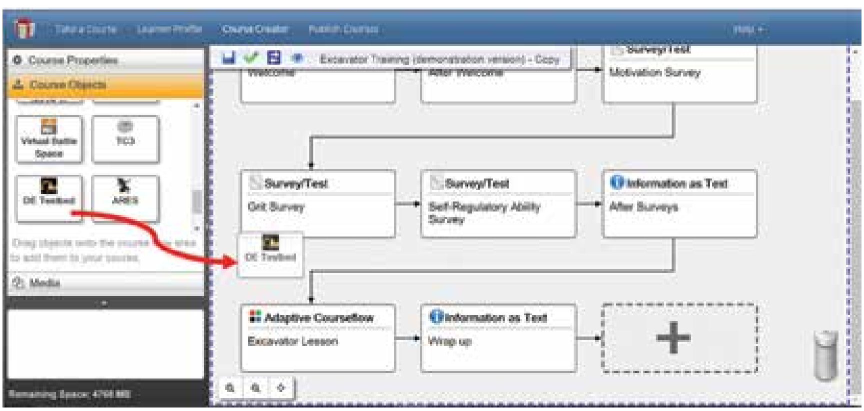

FIGURE 3
Drag and Drop Course Authoring in GIFT.

learner between four phases of instruction for a concept or set of concepts: rules, examples, recall, and practice. Figure 4 (right window) shows four concepts and associated rule phase content for an excavator simulator. In the rules phase, the learner is presented with terms and facts about the primary components of an excavator and their functions.

2.1.2 Reducing Cognitive Load by Eliminating Steps in the Authoring Process

One difficulty that was reported soon after we integrated GIFT with an external environment was that the authoring process was now significantly longer. For example, we integrated a game-based environment, Virtual BattleSpace (VBS), and authors reported that they now had authoring tasks in both GIFT and VBS. In VBS, scenarios would be defined using the VBS scenario editor. In GIFT, measures from VBS that were used to assess progress toward learning objectives had to be defined in GIFT as conditional classes so GIFT understood what they meant. For example, GIFT would define what elements constituted a location in VBS. However, each time you wanted to identify a location in VBS, it would require jumping from GIFT to VBS to identify that set of coordinates.

We decided to eliminate the extra steps in the GIFT authoring process by making a GIFT authoring window available in the native external environment. We call this GIFT Wrap (Figure 5) and it has been demonstrated for both VBS and the ARES tactical sandtable. We are attempting to generalize this interface for use in a larger array of systems by specifying the GIFT Wrap interface in much the same way we specified the GIFT Gateway.

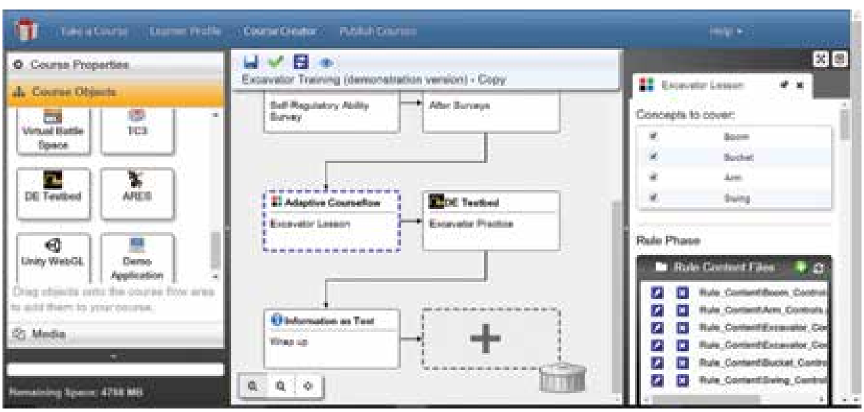

FIGURE 4
Configuring an Adaptive Courseflow Object in GIFT.

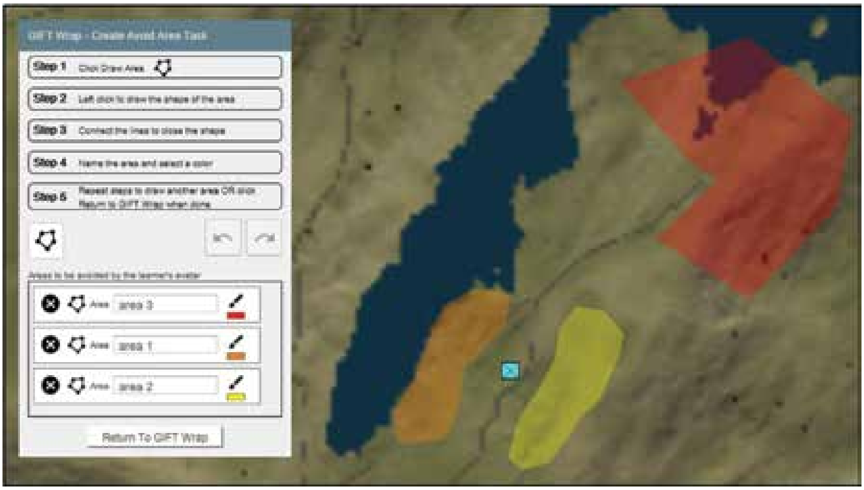

FIGURE 5
Configuring a Bounded Area in an External Environment using GIFT Wrap.

2.2 Automating GIFT Authoring Processes

Automating parts of the authoring process has been explored through different approaches with varying degrees of success including: automated content development from text sources and the development of wizards to guide inexperienced authors. Much of Army training differs from traditional ITS content (e.g., problem-based mathematics and physics tutors) in that it often requires conceptual knowledge (why you are doing something) in addition to procedural knowledge (what to do). We are seeking new methods to reduce the skill and time required to author scenario-based simulations and serious games to allow GIFT to automatically author variants of existing training scenarios which are relevant to the authors defined learning objectives.

2.2.1 Automated Scenario Generation

The method is called automated scenario generation (ASG; Zook et al, 2012) or evolutionary scenario generation (ESG; Luo, Yin, Cai, Zhong & Lees, 2016). This method focuses on how to use information from a "parent" scenario to generate hundreds or thousands of "child" scenarios and then rank order the child scenarios according to their relevance to a set of author-defined learning objectives. GIFT already allows authors to explicitly specify learning objectives known as "concepts". Additional detail on how GIFT functions can be found in the GIFT software documentation at www.GIFTtutoring.org.

The automated scenario generation method described would allow a GIFT-based tutor to customize (e.g., change difficulty level of the scenario) in real-time based on the learner's states (e.g., performance or emotion) or traits (e.g., personality) to optimize their learning, retention, and transfer of skills from training to the operational or work environment. This method would allow ITS developers who want to integrate GIFT with training simulation or serious games (e.g., Virtual Battle Space) to expand existing training capabilities to facilitate adaptive instruction with minimal additional burden on the scenario author.

2.2.2 Automated Concept Mapping

Another method to automate steps in the authoring process uses text-based documents (e.g., manuals, text books, course material, and web content) to generate hierarchical models of both expert measures and domain content (Kelsey, Ray, Brown, and Robson, 2015). These hierarchical models can then be used to generate the unique part of a GIFT-based tutor, the domain knowledge which includes content, question banks, assessments, and tutor interventions. While the hierarchical model generation shows promise, the error rate in the output (e.g., dialogue-based tutor or chatbot) remains too high for practical application at this time.

We will close our discussion of authoring challenges by touching on evaluation methods to compare various authoring tools and methods. This is difficult at best given the lack of standards between authoring systems and their resulting ITSs. Sottilare & Ososky (2017) developed an algorithm for measuring the complexity of GIFT-based tutors by assessing the complexity of the networks needed to define the complexity of their constituent learning concepts. Moving forward, they plan to expand the methodology within GIFT and examine methods to directly compare disparate tutors created by other authoring systems (e.g., Cognitive Tutor, AutoTutor).

3 GOAL #2: DEVELOPING EFFECTIVE INSTRUCTIONAL DECISIONS

Next, we examine the process of tutoring and the effectiveness of instructional decisions made by AISs through the lens of Merrill's Component Display Theory (CDT; 1983) and principles of personalized instruction.

3.1 Theory-Driven Instruction

As noted above there are four phases in CDT: rules, examples, recall, and practice. Rules are facts and principles associated with the domain of instruction.

As part of the rules phase in the domain of baseball, you would need to understand the concept of "shortstop". The examples phase provides models of successful process or behaviors where the learner is presented with examples of successful behaviors which in baseball would include demonstrations of how to hold a bat and position yourself in the batter's box. Next, the AIS would assess the recall of the learner about essential rules and examples. Finally, the learner would be placed in a training environment in order to practice and develop/maintain critical skills. In GIFT, we are developing gateways to allow the acquisition and assessment of data from a both live (learner in a real environment) and simulated practice (learner in a virtual environment.

GIFT provides three general actions by the tutor: instructional strategies, tactics, and policies. Strategies are recommendations by the tutor based on learner states and traits and are domain-independent. Strategies have been derived from the instructional and learning sciences literature. A meta-analyis provided their effect size and relation to learner attributes. Strategies are administered by GIFT's engine for managing adaptive pedagogy (eMAP; Goldberg, et al, 2012). EMAP recommendations include generalized plans of action or next steps by the tutor. Examples include "prompt for more information", "initiate a reflective dialogue", and "skip content based on prior knowledge."

Once a strategy is recommended by eMAP, it forwarded to the domain module where the tutor takes the recommendation and chooses and executes a tactic, an action with a domain-specific context. For example, a recommendation of "ask the learner a question" based on a assessed state of confusion results in a tactic selection of a specific question "what are the four principles of marksmanship?" for our marksmanship example. Policies have not yet been implemented in GIFT, but would be considered rules to be enforced by software-based agents to insure effective instruction. Examples of good instructional practices include mastery learning and error-sensitive feedback. Mastery learning is a policy of holding learners in a lesson until mastery of the concepts associated with that lesson have been demonstrated. Error-sensitive feedback weighs the criticality of learner mistakes to determine if and how often to provide corrective feedback.

3.2 Personalized Learning and Adaptive Instruction

The key to adaptive instruction is the personalization or tailoring of the tutoring process for each individual and team based on their learning needs and their preferences, but the critical element is the effect of the tutoring experience on learning. For this discussion, we extend the concept of *personalizing education* described by Woolf (2010) to be *personalizing all learning experiences* by including all forms of formal and informal education and training. To personalize learning, the tutor (human or machine-based) requires some level of understanding or a model of each individual learner or each group of learners during collective instruction (e.g., collaborative learning, team/group instruction).

This learner model (or the team model) must be updated to reflect changes in knowledge, skill, attitudes, beliefs, desires/goals, intentions, preferences, traits, and other states (e.g. emotions, performance). To maintain this learner model data, it must first be acquired (sensed, self- reported, retrieved, derived) and then classified into states (more transient) or traits (less transient) via an assessment process. This assessment of student learning posed by Woolf (2010) is particularly important due to the accuracy required to realize a significant learning effect (Sottilare, 2012; Sottilare, Ragusa, Hoffman, and Goldberg, 2013; Sottilare, 2014). Large inaccuracies in assessment of learner states will likely result in poor instructional decisions by the tutor. Once the learners' states and traits are classified, then the tutor must intelligently select an optimal course of action by the tutor (e.g., provide feedback, direct the learner to some activity for remediation) and then the cycle starts anew.

This tutoring process is described in the learning effect model (LEM) first proposed by Sottilare (2012) and later extended to include measures of effectiveness for both individuals and teams (Sottilare et al, 2017). The LEM forms the basis of the instructional process used by GIFT. Figure 6 highlights the interaction between the LEM for an individual learner (top) who may be part of a team (bottom). For the individual learner, GIFT may derive the level of domain competency from the Long Term Learner Model (LTLM; See Goal #3), a pre-test, or recent performance.

While the Team LEM has not yet been fully implemented in GIFT, it does provide a working model for its design. For teams, GIFT uses the LEM to assess both teamwork and taskwork. Teamwork examines the interactions of the team to assess attitudes, behaviors, and cognition in order to understand various antecedent states critical to team learning and performance (Sottilare et al, 2017). Teamwork is largely domain-independent. Whereas, team taskwork is domain-dependent in that its measures are dependent on the domain under instruction. In the Team LEM (Figure 6), GIFT assesses the progress of the team and the contributions of individuals to those team taskwork goals in order to determine its future strategy selections, recommendations, and tactics. Policies for teamwork and taskwork are still evolving.

The major challenge in the LEM is how to optimize decisions in an environment with a very large number of variables. Not only does each learner have a large number of states and traits that are either antecedents or direct indicators of learning, but the context in which that learning occurs is defined by many

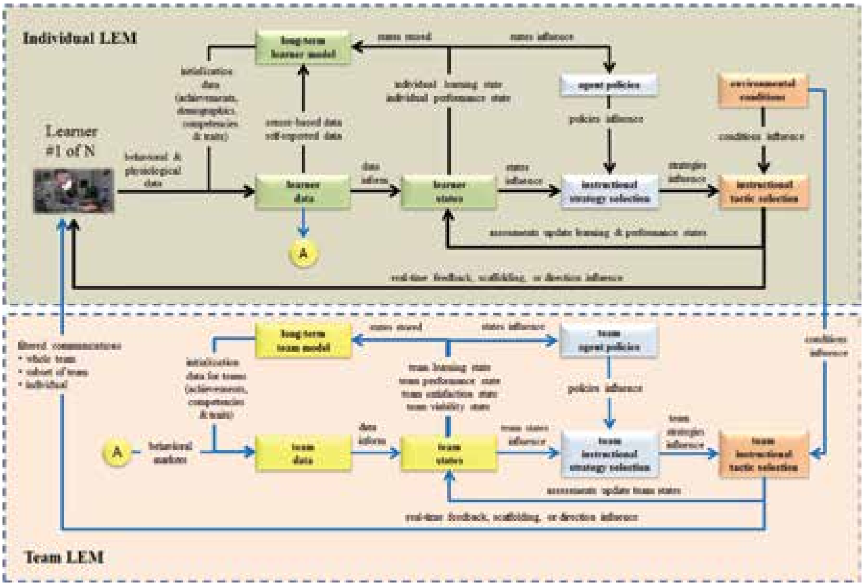

FIGURE 6
Adapting Team Instruction using the Learning Effect Model (LEM) for Individuals (top) and Team (bottom).

conditions in the environment (e.g., location, sequence in the course, ambient light, simulation resolution). These conditions may or may not affect learning or task performance depending on the domain under instruction.

GIFT helps to simplify this complex problem of evaluating effectiveness through two architectural principles: 1) instructional strategies (plans for tutor actions) are selected based only on learner states/traits and are independent of the domain of instruction and thereby independent of context; 2) selection of instructional tactics (actions executed by the tutor) are limited by the instructional strategy selection and then further limited by the context of the instruction. For example, if the learner's state is classified by a sensor measuring facial marker distrances as "confused", an appropriate instructional strategy selected by a GIFT-based tutor might be to ask the learner a question to assess their knowledge of a concept under instruction. An appropriate instructional tactic might be to select a specific question from a question bank that has been tagged as "medium difficulty" based on previous content presented and scenarios/problems discussed. This process is known as *instructional management* in GIFT.

Instructional management involves the automatic optimization of learning through the AIS's decisions and interactions during adaptive instruction. The tutor's goal is to enhance learning for that individual or team by adapting the instruction (e.g., changing the challenge level) based on the conditions of the learner and environment. Instructional management is the concept of automatically managing the delivery, pace, and sequencing of instruction including the assessment and response to changing states of the learner and affiliated instructional environments.

Goldberg, Sinatra, Sottilare, Moss, & Graesser (2015) documented instructional management goals and approaches for GIFT. A primary goal was to examine a variety of use cases in different task domains (e.g., cognitive, affective, psychomotor, and social) to understand the level of complexity relative to the conditions of the learner(s) and the environment and any competing outcomes (e.g., accelerated learning vs. retention). Understanding complexity aids the ability of the tutor to intelligently manipulate conditions to optimize outcomes.

One approach to managing complexity and uncertainty is to discover and develop modeling functions that account for uncertainty across various policies informing pedagogical decisions (e.g., content delivery, course navigation, and guidance). The objective here is to develop these functions to refine and optimize themselves through reinforcement learning mechanisms (e.g., Markov Decision Processes) over time as new interaction and performance data becomes available. A planning approach to quantify tutorial decisions and associated reward states has been prototyped (Rowe, Pokorny, Goldberg, Mott, and Lester, 2017) and is currently being validated through experimentation and will be incorporated into the GIFT cloud baseline (2018–1) in July 2018.

Another concurrent approach to addressing the optimization problem is based on observation of outcomes of human tutoring decisions. By observing the perception, judgment, and behaviors of expert human tutors to support practical, effective, and affordable learning experiences, we might be able to model their most effective strategies, tactics, and policies in software-based agents. One such example is modeling the scaffolding techniques of expert human tutors (Lepper, Drake, and O'Donnell-Johnson, 1997). Scaffolding is "an act of teaching that (i) supports the immediate construction of knowledge by the learner; and (ii) provides the basis for the future independent learning of the individual" (Holton and Clarke, 2006). Scaffolding involves "support, in the form of reminders and help, that the apprentice requires to approximate the execution of the entire composite of skills" (Blakeslee, 1997). As the learner grows in competence, support is proportionally withdrawn by the tutor until the learner is able to achieve the full learning expectations in the domain.

4 GOAL #3: MODELING LEARNER AND TEAM TRENDS AND COMPETENCY

This goal is focused on instructional methods for individual learners and teams which promote competency, skill or mastery in a domain. The development of a long term learner model (LTLM) is critical to understanding individual learner and team trends. Social learning is also discussed in terms of its relationship to developing shared mental models for collaborative learning or collaborative problem solving.

4.1 Long Term Learner Modeling of Competency

As noted earlier, a significant part of effective interaction between learners and tutors will depend upon the tutor's understanding of the learner which will be facilitated through the development of a LTLM. LTLMs are a record of learner data (e.g., knowledge, goals, experiences, achievements, preferences) over long enough periods to show trends of attitudes, behaviors, and cognition. Learner data can be used by AISs to infer/predict future learner states (e.g., domain competency of learners entering new instructional experiences) based on previous experiences. A primary challenge in building rapport/engagement with learners is accurately modeling their domain competence. Methods are needed to coalesce learner data into a quantifiable states representing varying levels of domain competence so the tutor may adapt appropriately to their learning needs.

The data in the LTLM may be persistent (little or no change over time) or transient (changes/decays over time). Examples of persistent data generally include gender, personality traits, and past achievements. Transient data includes knowledge and skill levels which decay overtime, emotional states, and interests. The importance and influence of specific learner attributes and their effect on competency and engagement will drive their presence in the LTLM. It is not unreasonable to see future LTLMs that contain fixed attributes (learner centric focus) and variable attributes which depend upon the domain of instruction. Recent designs for the LTLM include achievement statements based on the experience Application Program Interface (XAPI) to form a learner record store (LRS). Sottilare, Long, & Goldberg (2017) recommended changes to the XAPI standard to support modeling of learning, skill decay, and domain competency in order to provide a more comprehensive dataset for the LEM to act upon.

The LEM (discussed earlier in Goal #2) uses learner data to construct learner states or determine learner traits/preferences. Understanding these states/traits/

preferences enable the AIS to engage the learner with relevant content. For example, more experienced learners in a domain of instruction may anchor new learning to novice topics, but prefer more challenging content and problems. The ability of AISs to adapt learning experiences to be more germane to each individual means more engaging instruction in which the probability of learning increases.

4.2 Social Learning to Accelerate Competency

Another method of increasing learning is through peer or social learning. Woolf (2010) advocates the importance of supporting social learning with education technology and specifically calls out technology capabilities to "sustain continuous learning by active students in a way that enable students to communicate where they are located and to value learning". To build rapport between the learner and the tutor, AISs must enable social learning in way that learners can engage ITSs in the same or similar manner in which they engage human tutors and peer learners. In other words social learning methods should be enabled even when the interaction is only between a human learner and a machine-based tutor, but could also be applied to teams under instruction.

Social learning has been extensively covered in the team performance (Cannon-Bowers, Tannenbaum, Salas, and Volpe, 1995; Zachary et al, 1999; Cannon-Bowers & Bowers, 2011; Salas, 2015;) and computer-supported collaborative learning (CSCL) literature (Johnson & Johnson, 1986; Johnson & Johnson, 1999; Dillenbourg, 1999; Stevens, Berka & Sprang, 2009; Adamson, Dyke, Jang, and Rosé, 2014). Sottilare, et al (2017) distinguished team and collaborative learning models based on their goals and member behaviors and interaction. Teamwork models examine the behaviors and interaction of groups working toward common goals and are independent of the domain of instruction. Team taskwork models examine behaviors and interaction of groups working toward task objectives in a specific domain.

Finally, collaborative learning models represent member behaviors and interactions in aiding other members to reach learning objectives. Together these make up a class of collective instruction which is discussed further in Goal #5 of this article. Among the variety of instructional models (e.g., traditional classrooms, one-to-one tutoring, collaborative learning, and flipped classrooms), AISs may be a technology of choice to enable social learning inside and outside the classroom. Additional technology enablers of social learning include collaborative learning fora, and distributed/mobile learning solutions (see Goal #8).

5 GOAL #4: BUILDING RAPPORT AND ENGAGEMENT WITH LEARNERS

Another facet of social learning advocated by Woolf (2010) can be found in human tutors and AISs attempting to engage learners by tailoring content to their learning needs as identified by their states (e.g., prior domain knowledge), traits (e.g., personality trait of openness), and preferences (e.g., interests, learning style). AISs understand and model learners in order to guide them through an instructional experience. To duplicate the rapport developed between human tutors and learners, developers have integrated virtual humans (VH) in tutor interfaces. For example, one of the best known dialogue-based tutoring systems, AutoTutor (Graesser, Chipman, Haynes, and Olney, 2005), provides a VH interface to communicate feedback, support, and directions to the learner. GIFT also has VH capabilities as part of its tutor-user interface which may be driven via a conversation tree authored by the developer (fixed decision tree) or by calling AutoTutor conversational agents as a service (adaptive conversation). Research suggests that the physical characteristics of VHs influence the engagement and social presence of learners (Kim, Wei, Xu, Ko, & Ilieva, 2007), and learner engagement and decision-making can be also influenced by the perceived emotional expressions (both verbal and non-verbal) of VHs (Choi, Melo, Woo, and Gratch, 2012).

Evidence also implies that the channel of communication between the tutor and the learner or source modality (e.g., voice of unknown source, VH, or text) can make a difference in performance, retention and mental demand (Goldberg & Cannon-Bowers, 2015). Goldberg & Cannon-Bowers found that feedback from pedagogical agents in the form of VHs resulted in the largest retention outcomes during serious game play. They also found that feedback delivered as audio alone significantly lowered mental demand during game play.

VHs may also play a role in assessing learning, domain competency, or preferences through interactive dialogue with a learner or a team of learners. AutoTutor Lite (Hu, Cai, Han, Craig, Wang, and Graesser, 2009) defined a process for evaluating information provided by learners (verbal or written) based on two dichotomies: new/old information and relevant/not-relevant information. Using the tutoring process described by Person, Graesser, Kreuz, and Pomeroy (2003) where 1) the tutor asks a question or presents a problem to the learner, 2) the learner responds, 3) the tutor provides brief feedback, 4) the learner and the tutor collaboratively improve the quality of the answer, and 5) the tutor assesses the learner's understanding of the answer. During this tutoring process, the tutor tracks whether information provided by the learner is old or new and relevant or not relevant to the domain of instruction or the question posed by the tutor. Old or repeat information from the learner does not improve the quality of the answer, but new information may assuming it is relevant to the domain. This process of collaboratively refining an answer develops the trust of the learner and maintains high levels of engagement.

The current version of GIFT supports service calls to AutoTutor to enable interactive dialogue. GIFT provides context, AutoTutor manages the dialogue, and then returns control to GIFT. Experimentation with GIFT and AutoTutor indicate that tighter coupling is required between the two functions in the way of datasharing to communicate more detailed results of interactive dialogue in AutoTutor, but a basic capability to support interactive dialogue exists.

Another option to support interactive dialogue in GIFT is to drive dialogue through a commercial or academic VH software package. GIFT has an embedded Media Semantics Character that can provide feedback based on assessments in GIFT or be driven through a conversational tree in GIFT (Figure 7).

Social aspects of human-VH dialogue may also be reinforced by agent-based VHs using human information processor models like GOMS (goals, operators,

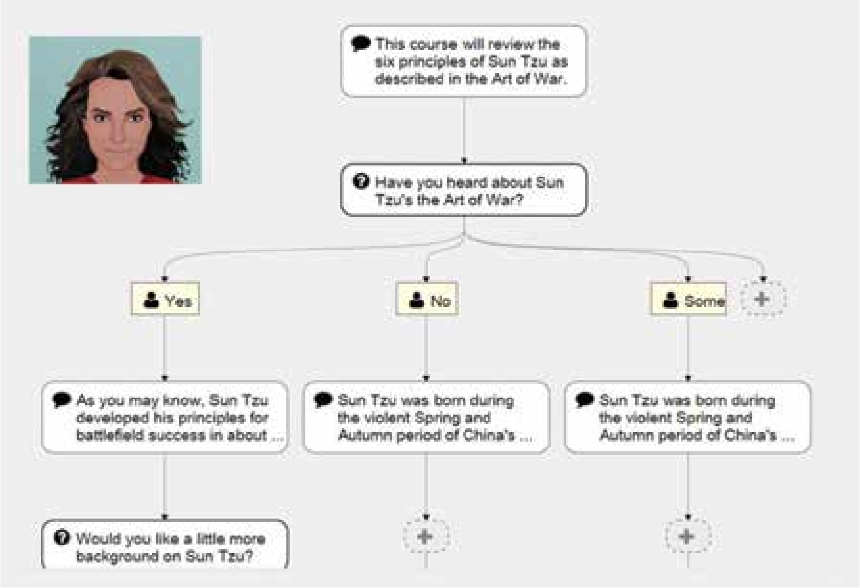

FIGURE 7
A Simple Conversational Tree in GIFT.

methods, and selectors; John & Kieras, 1996) to guide interaction between humans and agents. A set of goals (e.g., insure mastery of concept B, demonstrate skill C) are defined for the agent to achieve with the learner. A set of operators or possible actions to be taken by the tutoring agent are defined based on the domain context. Operators may be perceptual, motor or cognitive acts that are used by the agent to influence any aspect of the learner's state (e.g., cognitive) or the task environment (see Figure 1). Methods are a set of steps or actions a procedure used by the agent to accomplish an assigned goal. Finally, selectors are rules used by the agent to select the most appropriate method available to reach an assigned goal or to deconflict competing goals (e.g., Goal A is a higher priority than Goal B).

The Virtual Human Toolkit (VHT; Gratch, Hartholt, Dehghani, and Marsella, 2013) takes a more sophisticated approach to VHs as tutoring agents. The assessment and decision-making functions of VHs generated by the VHT have their basis in psychological theories of emotion (Gratch & Marsella, 2005), natural language understanding (Traum, 2008), and human cognition (Swartout et al, 2006). The evolution of GIFT as a multi-agent architecture is being driven by the need to describe GIFT functions as services, and the VHT will be one set of services that allow authors to generate agents to support various interactions between GIFT-based tutors and learner populations.

Another aspect of how VHs or agents can support the development/maintenance of rapport with learners is illustrated in the concept of teachable agents like "Betty's Brain" (Leelawong and Biswas, 2008). The concept of learning by teaching others is a powerful methodology in which computer-based, domain-independent teachable agents are used to motivate learners to learn more so they can pass on this learning to the agent. Other concepts for teachable agents are represented in tutoring methodologies like s where human learners interact with a virtual teacher and a virtual learner managed by intelligent agents (Cai, Feng, Baer & Graesser, 2014). Learners that might hesitate to interact with the virtual teacher may be more comfortable interacting with the virtual peer. As noted above, the ability to configure a virtual peer to take on the characteristics of the learner's culture and preferences may enhance the engagement of the learner and thereby improve learning, performance, and retention.

The VH literature suggests that AISs can influence learning and performance through the configuration or manipulation of physical attributes and interactions of a VH to engage the learner in the instructional experience. Future versions of the GIFT authoring tools will need to incorporate elements to allow AIS authors to configure the physical aspects of VHs and adapt their culture and verbal responses based on learner preferences.

6 GOAL #5: MODELING COLLECTIVE INSTRUCTIONAL DOMAINS

Most tutors today are designed for individual learners, but a growing need has been identified to apply AISs to collective training and education. The move from individual to collective instruction will increase the complexity of most domains, increase authoring requriements, and increase the complexity of the assessment and intervention process for the tutor. As noted in our Goal #2 discussion, collective or team domains are not a simple multiple of the number of team members in a domain. There is additional complexity in the interactions between team members, their roles and responsibilities, and the leadership present in the group. All these factors (and more) affect the learning and performance of groups.

Given the complexity and effort required to author collective tutors, what is the motivation to build them and how do we want the AIS to interact with learners in a team? Similar to individual tutors, the motivation to build collective or team tutors is to efficiently and effectively guide instruction in the absence of human tutors. Since many tasks are of a collective nature, it natural to want to apply the effectiveness of ITSs to teams, the foundational element of many organizations. The challenges associated with collective domains may seem large now, but successful models are developing in collaborative learning and small team taskwork.

One challenge is to understand the modeling of groups with different objectives: teamwork, team taskwork, and collaborative learning. Teamwork is focused on the functions of the team independent of task domain. A model of teamwork derived from Burke, Stagle, Salas, Pierce and Kendall (2006) examines the influence of various team functions:

- Communication – disclosing information to or exchanging information between team members
- Conflict – processes within the team to recognize and resolve conflict
- Coaching – all leadership activities required to maintain a well functioning team
- Cooperation – motivational drivers within the team to achieve the team's goals
- Coordination – behavioral mechanisms within the team used to accomplish goals
- Cognition – common understanding of goals, roles, responsibilities, and domain; sometimes referred to as shared mental models
- Context – team norms (rules, best practices) and interaction with the domain knowledge (content)

This model was used to understand various team states and examine antecedents of team learning and performance in the literature from 2003 to 2013 (Sottilare et al, 2015; Sottilare et al, 2017). Specifically, this large meta-analysis was conducted to complement and extend several meta-analyses published during the early 2000s (e.g., leadership by Burke et al., 2006; cohesion by Beal et al., 2003; team conflict by DeDreu & Weingart, 2003). Antecedents of team performance included team states and their affiliated behavior markers: cohesion, collective efficacy, communication and leadership. While antecedents of team learning included: trust, cohesion, and conflict management. The remaining challenge is to implement assessment methods within GIFT to classify these antecedent team states via the recognition of key behaviors among team members. Once this is done, interventions can be constructed for the tutor and a validation of the whole team tutoring model will be undertaken.

While teamwork is about the examination of team interactions, and efficient/effective communication across the team, team taskwork is about training to learn how to do a specific task as a team. In order to understand the team's progress toward learning objectives (e.g., mastering tasks, recalling procedures) the AIS must be able to measure and interpret specific actions by each team member and understand its relationship to their roles and responsibilities. It is complex enough to acquire individual learner data and use that to infer individual changes in their individual performance state. Tutoring teams to perform taskwork in specific domains also involves understanding how teamwork and collective measures of team task achievement rollup to a team mastery measurement.

The major challenge for team taskwork instruction is that increases in task complexity, team interdependence, and team dynamics combined with decreases in task definition make team measures and assessment more difficult (Sottilare & Ososky, 2017). Task complexity is measured in GIFT as the number of leaf nodes in a tree of concepts or learning objectives. In other words, the number of terminal concepts that must be measured for assessment of a team task. The larger the number of leaf nodes and dependencies between elements of the task, the more complex the task and more difficult it is to assess. Differences in member competence in tasks can also diminish the ability of the team to perform at optimal levels. If there are five concepts to be mastered for a given team task (Task 1) and for three of those concepts, Team Member A is dependent on the skill and output of Team Member B, then there is a high likelihood that this task is more complex than a team task with one interdependent concept, and that more complex tasks are more difficult to master.

Referring back to the LEM, (Figure 8), it is easy to see that errors in acquiring learner data (e.g., no input or failed sensors) can lead to misclassification of a

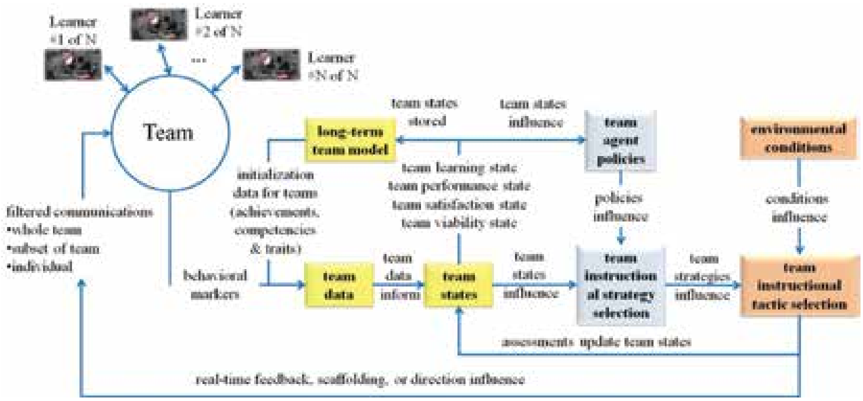

FIGURE 8
A Team Learning Effect Model (LEM; Sottilare et al, 2017).

higher percentage of learner states and therefore lead to a higher error rate in classifying team states. Misdiagnosis of team states lead to higher error rates in both instructional strategy and tactical decisions.

Collaborative learning environments may be treated similarly to team taskwork environments in that they both seek to educate/train groups in a particular domain, but differ slightly in their goals. Collaborative learning is about learning to do something that can be performed by an individual, but is learned through a group process in which each group member can contribute to the learning of other group members. In team taskwork environments, it is often assumed that individuals possess the skills required to perform their assigned roles. While this may not always be true, the focus of team taskwork training is to learn how to work as a team to perform a task that is usually not performed by an individual. For example, a collaborative group might share techniques for solving a quadratic equation, but it is not a team task. Conversely, someone who might have collaboratively acquired knowledge of all the positions on a basketball team would not go out and perform alone. Teamwork is not taskwork, but teamwork includes collaboration required to perform collective tasks. Taskwork is not necessarily collaborative learning, but may include collaboration. While these tasks are different in some ways, they are not mutually exclusive of one another.

A major challenge is the ability for a tutoring architecture like GIFT to represent all types of collective tasks including methods to acquire individual learner data, assess individual and team states, and then optimally select the most appropriate policy, strategy or tactic. Building reliable models usually means collecting lots and lots of data. This may be practical in large organizations where there are high density courses/lots of learners, but may not be practical where assessing team states based on a large number of conditions appear for the learners and the instructional environment. This means many sets of conditions may be rarely seen even with large datasets. Seeking methods to reduce the amount of data needed to build dependable models should be a significant goal for the AIS community in the nearterm.

7 GOAL #6: EXPANDING ADAPTIVE INSTRUCTION TO A BROADER ARRAY OF TASK DOMAINS

As noted earlier, ITSs have been prominently tied to well-defined, cognitive domains like mathematics, physics, and software programming. In these domains, ITS have demonstrated significant effect on the learning and performance of individuals, but many instructional domains involve physical tasks and many more involve collectives or teams. ITSs have not yet demonstrated their effectiveness in these domains. In order to show the influence on learning effectiveness, learning efficiency (accelerated learning), retention, transfer of skills, and relevant team outcomes (e.g., team learning, performance, satisfaction, and viability), ITSs and their extended counterparts AISs must be able to produce prototype environments with adequate methods of measurement and assessment. Murray (1999, 2003) noted the need for a rapid ITS prototyping capability to be able to evaluate various ITS designs and this includes new domains in AISs.

With a goal of expanding adaptive instruction to a broader array of task domains, a set of questions remain open and are discussed below:

- How do we represent new domains to learners so learning takes place?
- What methods are needed to acquire data associated with identified measures?
- What measures are needed to assess learning of individuals and performance of teams?
- Are there classes of task domains that have common factors so that they may be treated the same or similarly during authoring, instructional delivery, or evaluation processes?

7.1 Representing New Domains

In the nearterm, extending ITSs to new domains may be limited to applying techniques used in cognitive (decision-making and problem solving) domains to case studies in affective domains (e.g., emotional intelligence, decision-making based on moral judgment). The basic functionality of ITSs would not have to change much in order to represent affective case studies, but to make the leap to psychomotor domains and social/team/collective instructional domains requires some fundamental change to how ITSs are structured and how they function.

Psychomotor tasks involve the enhancement of physical movement, coordination, and the use of the motor-skills through deliberate practice and is measured in terms of speed, precision, distance, procedures, or application of techniques (Simpson, 1972). The acquisition of psychomotor task measures is more complicated than for task domains in which the learner's inputs are captured directly by the computer. Significant challenges exist in understanding the learner's interaction within real/live environments. Vargas-González, Williamson, LaViola & Sottilare (2017) have anticipated the need to model this interaction between the learner, the real environment, and perhaps even virtual augmentations within the real environment (augmented reality - AR). The combination of see-through displays (e.g., Microsoft Hololens) with multi-modal interaction techniques show high potential as AISs for instructing psychomotor tasks when linked to ITS technologies. To support this integration, a plugin or interop must be developed to allow the external applications/environments to communicate within GIFT. In Vargas-González's application, a new interop was developed to link Unity 3D and a XML Remote Procedure Call (XMLRPC) server client to communicate between GIFT and a Windows Universal Platform Applications (HoloLens). Recently, a protoype user interface for authoring interoperable augmented reality environments has been implemented in a GIFT software development branch and is being enhanced to improve it usability.

Specific tasks are being examined to understand what is needed to assess learning and performance in new domains. Laviola et al (2015) examined aspects of tutoring tasks in augmented reality environments in which instruction could be provided almost anywhere and which led to authoring application developed by Vargas-González and colleagues (2017). In another psychomotor task, Sottilare, Hackett, Pike, and LaViola (2016) hypothesized the use of pressure sensors and smart glasses to train hemhorrage control tasks. Pressure sensors placed in tourniquets and pressure bandages would be used to measure the learner's ability to reduce bleeding.

7.2 Acquiring Data and Assessing Learning in New Domains

One approach to assessment of learning and performance in a variety of task domains is evidence-centered design (ECD; Mislevy & Haertel, 2006), an approach to constructing educational assessments in terms of evidentiary arguments which provide clear and convincing proof of the learner's state(s). Usually measures of elements within a task are identified as assessment criteria and form the basis of modeling the learner's performance. These measures could be decisions or actions taken by the learner. ECD provides a process which links features of tasks to various learning or performance outcomes, and there by provide evidence of learning or performance. The ECD process is organized into five layers: domain analysis, domain modeling, conceptual framework, implementation, and delivery (Mislevy & Riconscente, 2006). In the domain analysis layer, information is mined from the task domain to determine how knowledge is developed or acquired by the learner, used by the learner, and shared with other learners. In the domain model, specific assessment arguments are developed in narrative form based on the domain analysis. In the conceptual framework, specific assessment arguments are formed into structures and specification for tasks, tests, evaluations, and measurement models. In the implementation layer, the assessment is produced and includes tasks and calibrated measurement models. Finally, in the delivery layer, this process coordinates the interactions of the learner and sequencing of tasks and also provides scoring and reporting.

Measures in the psychomotor domain are generally captured remotely (e.g., motion capture via remote sensing) or derived from other measures. This presents challenges with potential data gaps or receipt of corrupted data. A robust AIS design should consider the accuracy of these measures and the reliability/availability of sensor data for classification. GIFT currently provides a standard gateway specification which allows authors to develop interops to link sensors and external systems to GIFT and to receive and interpret data from GIFT.

7.3 Grouping Task Domains into Classes

The advantage of grouping task domains into classes is the ability to examine commonalities in data structures, measures, acquisition methods, and assessment techniques for each class in order to author, deliver, and evaluate adaptive instructional technology. Initial studies are focusing on defined task domain classes: cognitive (Bloom, 1956; Krathwohl, 2002), affective (Krathwohl, Bloom, & Masia, 1964), psychomotor (Simpson, 1972), and social (Soller, 2001). Sottilare & LaViola (2015) and Sinatra & Sottilare (2016) examined similarities in psychomotor tasks in attempting to categorize tasks beyond traditional desktop computer-based tutoring in traditional cognitive task domains (e.g., mathematics, reading comprehension, physics) and identified similarities in measures and data acquisition techniques. Sinatra & Sottilare's analysis is a two dimensional taxonomy of domains which evaluates tasks by complexity and definition resulting in three viable groupings of tasks: well-defined, low complexity; well-defined, high complexity; and ill-defined, high complexity. Since by their nature ill-defined domains are complex, there is no category for ill-defined, low complexity. Complexity is defined by the structure of the domain knowledge and the number of concepts to be assessed (Sottilare & Ososky, 2017).

Given the variety of domains which might be adaptively instructed, a couple of key questions arise when discussing the potential of ITSs and AISs into new task domains:

- Will AISs be as effective in new domains as they have been in traditional domains?
- Are different methods needed to determine the effectiveness and efficiency of adaptive instruction in various domains?

This topic will be addressed in Goal #7 below.

8 GOAL #7: EVALUATING THE EFFECTIVENESS AND EFFICIENCY OF ADAPTIVE INSTRUCTIONAL SYSTEMS

A key part of Woolf's (2010) personalizing learning goal is the ability of the AIS or ITS to assess learning resulting from personalization and involves the need for methods to evaluate the effectiveness and efficiency of adaptive instructional technologies (tools and methods). The goals of enhanced effectiveness and efficiency differ, but are complementary. For effectiveness, we seek to improve the capacity of the learner in a fixed period of time, and for efficiency, we seek to accelerate learning or reduce the time to learn a fixed amount of material (Sottilare, 2017a). It is worth noting that training and education goals differ (Fletcher, 2017). Training objectives are focused on learning to do a specific task or set of tasks so the learner can do these tasks in the operational or work environment. Educational objectives are much broader and focused on preparing the learner to perform in yet unknown work environments. These differences in goals may mean the instructional approaches should also differ, and then so should the methods to evaluate their effectiveness/efficiency.

Given the complexity of AISs based on the number of conditions represented in the learner(s) and the environment, and the large degrees of freedom represented by the instructional decision space, it is often difficult to just look at an instructional situation and apriori understand what should be done to optimize learning or retention or performance or transfer of skills. As discussed in Goal #2, a large number of studies have been reviewed as part of meta-analyses to initialize best practices for AISs, but still must be validated through experimentation.

To this end, it is critical to make big data available to reinforcement machine learning processes to understand adaptive instructional decisions and the resulting value or effect. Over time, the evaluation of these decisions will result in improved effectiveness.

AISs must be self-evaluating and self-regulated or allow for rapid analysis by other systems by allowing them access to run-time data. There are many tools that could be used to examine effectiveness, but we will mention a few open and commercial technologies for reference. Each of these tools may be used independently or in combination. For example, pre-processed (filtered) data from RapidMiner might be used by GIFT to classify learner states. Because GIFT is a modular architecture, it is possible to integrate nearly any open source big data analysis tool. The inclusion of these tools in this list does should not be construed as an endorsement, but merely a few easily accessible tools known by the author to examine questions of AIS effectiveness.

First, the GIFT testbed (Figure 2) combines common elements of ITSs, the learner model, domain model and instructional model to form and experimental system which can be evaluated against desired outcomes: learning, performance, retention, or transfer of skills. The testbed and its experimental system can also be used to conduct comparative analyses: AIS vs. traditional training methods, intervention vs. non-intervention strategies, the relative importance of new learner model attributes or new policies, strategies or tactics.

An evolving data repository with analytic capabilities is LearnSphere (Koedinger, Liu, Stamper, Thille & Pavlik, 2017). LearnSphere is decendent of DataShop (VanLehn et al, 2007). Funded by the National Science Foundation, LearnSphere stores educational data associated with ITSs, AISs, educational games and massively open online courses (MOOCs) so course developers and instructors will be able to improve adaptive instruction through data-driven, evidence-centered course design. Unlike DataShop which was located on a single server, LearnSphere can support many data repositories or spheres in the cloud with varying degrees of sharing.

Another open source analysis tool is the Waikato Environment for Knowledge Analysis (WEKA; Eibe, Hall & Witten, 2016) which is a suite of machine learning algorithms and tools written in Java and provided as free software under the GNU General Public License. While WEKA has generally been an offline tool, the public availability of its software allows developers to apply WEKA to their systems so it could be used for real-time evaluations.

Free commercial tools for the evaluation of big datasets are numerous. We won't try to mention them all here, but only reference a few that have been applied to datasets and documented in the instructional literature:

- RapidMiner (Mierswa, Wurst, Klinkenberg, Scholz & Euler, 2006)
- Orange (Demšar, Zupan, Leban, & Curk, 2004)
- Knowledge Extraction based on Evolutionary Learning (KEEL; Alcalá, Garcıa, del Jesus, Ventura, & Garrell, 2007)

9 GOAL #8: SUPPORTING DISTRIBUTED/MOBILE LEARNING

A critical path to Woolf's (2010) goals of personalizing learning and diminishing barriers are tied to the flexibility of the AIS to provide learning opportunities in a variety of settings, and a significant challenge to making AISs practical for widespread use is the ability to apply adaptive instructional principles at a distance. To improve accessibility and usability, GIFT is available as a cloud-based AIS via Amazon Web Services (AWS). The advent of team modeling in GIFT will allow distributed team learning via CloudGIFT. We are also beginning to examine opportunities to leverage mobile platforms (smartphones and tablets) as shown in Figure 9, but there are some significant challenges in assessing learner states based only on sensors available in a mobile platform (e.g., GPS, camera) and design decisions need to be made about how to handle learner data when there is poor or no connectivity. There are options for storing and transmitting data when connectivity is available again, but storage may be problem when datasets are large.

In laboratory or classroom settings, it is possible to unobtrusively collect information about the learner through sensor suites and self-report data. The use of sensors at a distance is a primary challenge. For example, a learner is on the move and has a mobile device through which he will receive instruction. While smartphones have a bevy of sensor to report location/position and some behaviors, they are just beginning to be able to capture physiological data reliably as they are paired with smart watches (e.g., Samsung, Apple, or Google) and other sensors. The limitation to these technologies now are the lack of processing power onboard the mobile device to assess complex states in real-time based on data streams. Presently, it is impractical to send streams of physiological data to a central server for processing. Smartglass manufacturers also found this out early and began using offloading some data and calculations to the learner's smartphone with limited success. This problem becomes more difficult as we scale up from individual learners to teams. Capturing some behavioral markers necessary to classify teamwork or taskwork states of teams is currently not practical in mobile learning environments and may not be practical.

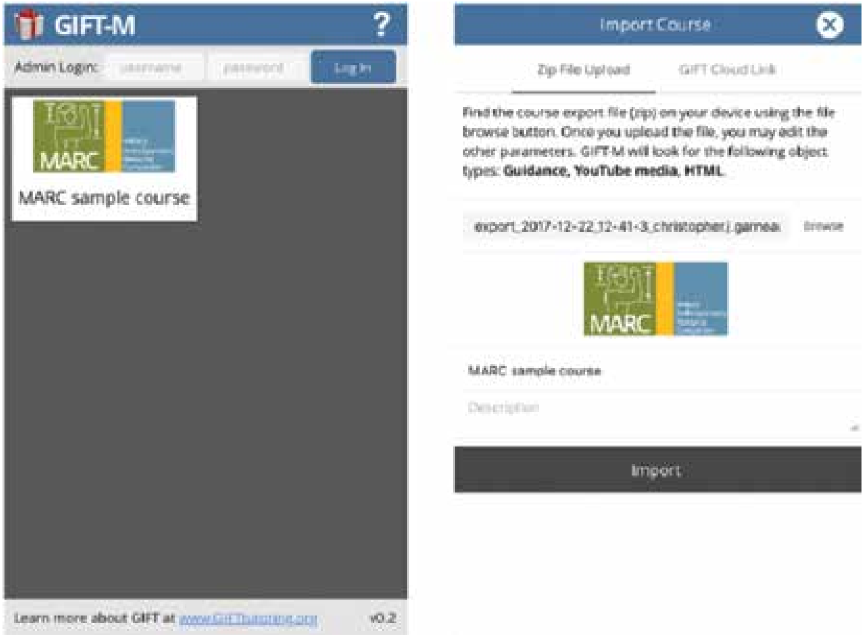

FIGURE 9
A Mobile Concept for GIFT.

In the nearterm, we may have to realize the limitations of the technologies and design instruction around those limitations. Longer term, we want to solve these problems. If we are fortunate, commercial technology will come and solve some or all of these problems. While we are waiting, it would be wise to identify alternate methods to be able to conduct the same assessments. If mobile sensors on the instructional platform (e.g., smartphone, tablet, laptop) are inadequate, we might devise methods to use more reliable/available data sources automatically.

10 A FINAL WORD

We close out the discussion of design goals, challenges, and emerging solutions by noting that while the comprehensive review presented in this paper covers a myriad of topics, it is far from exhaustive. This article serves as a companion document to expand on the topics covered in our AERA 2017 talk (Sottilare, 2017b). Our goal here was to present a state of art and practice in ITS design in order to engage the learning research and technology community in the pursuit of

TABLE 1
Goals and Strategies for the Design of Adaptive Instructional Systems (AISs).

Goal #1	Goal #2	Goal #3	Goal #4	Goal #5	Goal #6	Goal #7	Goal #8
Developing Efficient Authoring Processes	Developing Effective Instructional Decisions	Modeling Individual Learners and Teams	Building Rapport and Engagement with Learners	Modeling Collective Instructional Domains	Expanding Adaptive Instruction to a Broader Array of Task Domains	Evaluating the Effectiveness and Efficiency of AISs	Supporting Distributed/ Mobile Learning
Improve Usability by Reducing Cognitive Load and Extraneous Information	Adopt Learner-Centric, Domain-Independent Strategies & Recommendations	Adopt Long-Term Learner and Team Models to Classify/ Predict Long Term States (e.g., competency)	Integrate Virtual Humans with Natural Language Interaction (e.g., AutoTutor, Virtual Human Toolkit)	Assess Teamwork (domain-independent interaction)	Develop Expert Models Assess and Intervene During Psychomotor Tasks	Model the Complexity of ITSs for Objective Authoring Efficiency Comparisons	Develop Assessment Methods for Distributed and Mobile Learning Domains
Automate Scenario Generation to Extend Content for Adaptive Instruction	Bound Tactics (tutor actions) by Selected Strategies & Policies	Competency Modeling Supported by Experience API (XAPI) Statements of Achievement	Speed up Authoring using Virtual Humans and Conversation Trees for Bounded Conversations	Assess Team Taskwork (domain-dependent measures)	Develop Methods to Assess and Intervene During Social/Team/ Collaborative Tasks	Use GIFT Testbed Methodology to Evaluate Influence of Various Models and Methods	Develop Methods to Manage Intermittent Connectivity during Mobile Learning
Automate Content Curation	Manage Real-Time Instruction Using Adaptive Courseflow in GIFT (based on Merrill's Component Display Theory	Support Social Learning to Accelerate Knowledge & Skill Acquisition	Develop Long Term Learner and Team Models to Identify Trends and Preferences and Enhance Engagement	Build Reliable Models to Drive AIS interventions	Examine Methods to Enhance Collaborative Problem Solving as a Domain of Instruction	Use Open Source Tools for Data Analysis & LearnSphere to Share Data	Augment Reality to Support Tutoring in the Wild (beyond the desktop)

identified goals. A number of very real challenges are yet to be overcome to realize fully functional AISs that can support authoring, instruction, and evaluation of nearly any task domain and do it efficiently and effectively. This article identifies a few of the many steps required to realize that goal. Below is a table which summarizes many of the strategies and recommendations made in this article.

REFERENCES

Adamson, D., Dyke, G., Jang, H., & Rosé, C. P. (2014). Towards an agile approach to adapting dynamic collaboration support to student needs. *International Journal of Artificial Intelligence in Education*, 24(1), 92–124.

Alcalá, J., Sánchez, L., Garcıa, S., del Jesus, M. J., Ventura, S., & Garrell, J. M. (2007). KEEL: A data mining software tool for assessing the performance of knowledge extraction-based on evolutionary algorithms. *Soft computing*.

Beal, D. J., Cohen, R. R., Burke, M. J. & McLendon, C L. (2003). Cohesion and performance in groups: A meta-analytic clarification of construct relations. *Journal of Applied Psychology, 88*(6), 989–1004.

Blakeslee, A. M. (1997). Activity, context, interaction, and authority: Learning to write scientific papers in situ. *Journal of business and technical communication*, 11(2), 125–169.

Bloom, B. S. (1956). Taxonomy of educational objectives. Vol. 1: Cognitive domain. New York: McKay, 20–24.

Burke, C. S., Stagl, K. C., Salas, E., Pierce, L., & Kendall, D. (2006). Understanding team adaptation: A conceptual analysis and model. *Journal of Applied Psychology*, 91(6), 1189.

Cai, Z., Feng, S., Baer, W., & Graesser, A. (2014). Instructional Strategies in Trialogue-based Intelligent Tutoring Systems. Design recommendations for intelligent tutoring systems, 2, 225–235.

Cannon-Bowers, J. A. and Bowers, C. (2011). Team development and functioning. In S. Zedeck (Ed.) *APA handbook of industrial and organizational psychology, Vol 1: Building and developing the organization* (pp. 597–650). Washington, DC, US: American Psychological Association.

Cannon-Bowers, J. A., Tannenbaum, S. I., Salas, E., & Volpe, C. E. (1995). Defining competencies and establishing team training requirements. *Team effectiveness and decision making in organizations*, 333, 380.

Choi, A., Melo, C. D., Woo, W., & Gratch, J. (2012). Affective engagement to emotional facial expressions of embodied social agents in a decision-making game. *Computer Animation and Virtual Worlds*, 23(3–4), 331–342.

De Dreu, C. K., W., & Weingart, L. R. (2003). Task versus relationship conflict, team performance, and team member satisfaction: A meta-analysis. *Journal of Applied Psychology*, *88*(4), 741–749.

Demšar, J., Zupan, B., Leban, G., & Curk, T. (2004, September). Orange: From experimental machine learning to interactive data mining. In European Conference on *Principles of Data Mining and Knowledge Discovery* (pp. 537–539). Springer Berlin Heidelberg.

Dillenbourg, P. (1999). What do you mean by collaborative learning. *Collaborative-learning: Cognitive and computational approaches*, *1*, 1–15.

Eibe Frank, Mark A. Hall, and Ian H. Witten (2016). The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.

Fletcher, J. D. (2017). Intelligent Tutoring and a Role for Big Data. American Education Research Association (AERA) Conference, San Antonio, Texas.

Goldberg, B., Brawner, K. Sottilare, R., Tarr, R., Billings, D., & Malone, M. (2012). Use of Evidence-based Strategies to Expand Extensibility of Adaptive Tutoring Technologies. In Proceedings of the *Interservice/Industry Training Simulation & Education Conference*, Orlando, Florida, December 2012.

Goldberg, B., & Cannon-Bowers, J. (2015). Feedback source modality effects on training outcomes in a serious game: Pedagogical agents make a difference. *Computers in Human Behavior*, 52, 1–11.

Goldberg, B., Sinatra, A., Sottilare, R., Moss, J., and Graesser, A. (2015). Instructional Management for Adaptive Training and Education in Support of the US Army Learning Model: Research Outline. *Army Research Laboratory* (ARL-SR-0345), November 2015.

Graesser, A. C., Chipman, P., Haynes, B. C., & Olney, A. (2005). AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education*, 48(4), 612–618.

Gratch, J. & S., Marsella (2005). Lessons from Emotion Psychology for the Design of Lifelike Characters. Applied Artificial Intelligence, 19, pp 215–233.

Gratch, J., Hartholt, A., Dehghani, M., & Marsella, S. (2013). Virtual humans: a new toolkit for cognitive science research. *Applied Artificial Intelligence*, *19*, 215–233.

Hanks, S., Pollack, M. E. and Cohen, P. R. (1993). Benchmarks, Test Beds, Controlled Experimentation, and the Design of Agent Architectures. *AI Magazine* Volume 14 Number 4.

Holton, D., & Clarke, D. (2006). Scaffolding and metacognition. *International journal of mathematical education in science and technology*, *37*(2), 127–143.

Hu, X., Cai, Z., Han, L., Craig, S. D., Wang, T., & Graesser, A. C. (2009, July). AutoTutor Lite. In Proceedings of the 2009 conference on *Artificial Intelligence in Education*: Building learning systems that care: From knowledge representation to affective modelling (pp. 802–802). IOS Press.

John, B. E., & Kieras, D. E. (1996). The GOMS family of user interface analysis techniques: Comparison and contrast. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 3(4), 320–351.

Johnson, R. T., & Johnson, D. W. (1986). Cooperative learning in the science classroom. *Science and children*, *24*, 31–32.

Johnson, D. W., & Johnson, R. T. (1999). What Makes Cooperative Learning Work.

Kelsey E., Ray F., Brown D. and Robson R. (2015). Design of a domain-independent, interactive, dialogue-based tutor for use within the GIFT framework. In: Sottilare R, Sinatra A, editors. *Proceedings of the 3rd Annual GIFT Users Symposium (GIFTSym3)*; 2015 Jun 17–18; Orlando, FL. Army Research Laboratory (US). ISBN 978-0-9893923-8-9.

Kim, Y., Wei, Q., Xu, B., Ko, Y., & Ilieva, V. (2007). MathGirls: Toward Developing Girls' Positive Attitude and Self-Efficacy through Pedagogical Agents. *FRONTIERS IN ARTIFICIAL INTELLIGENCE AND APPLICATIONS*, *158*, 119.

Koedinger, K., Liu, R., Stamper, J., Thille, C., & Pavlik, P. (2017, March). Community based educational data repositories and analysis tools. In Proceedings of the *Seventh International Learning Analytics & Knowledge Conference* (pp. 524–525). ACM.

Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). Taxonomy of educational objectives, handbook ii: affective domain. New York: David McKay Company. Inc. ISBN 0-679-30210-7, 0-582-32385, 1.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into practice*, 41(4), 212–218.

LaViola, J., Sottilare, R., Garrity, P., Williamson, B., Brooks, C., and Veazanchin, S. (2015). Using Augmented Reality to Tutor Military Tasks in the Wild. In Proceedings of the *Interservice/Industry Training Simulation & Education Conference*, Orlando, Florida, December 2015.

Leelawong, K., & Biswas, G. (2008). Designing learning by teaching agents: The Betty's Brain system. *International Journal of Artificial Intelligence in Education*, 18(3), 181–208.

Lepper, M. R., Drake, M., & O'Donnell-Johnson, T. M. (1997). Scaffolding techniques of expert human tutors. In K. Hogan & M. Pressley (Eds), *Scaffolding student learning: Instructional approaches and issues* (pp. 108–144). New York: Brookline Books.

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*.

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

Mierswa, I., Wurst, M., Klinkenberg, R., Scholz, M., & Euler, T. (2006). YALE (now: RAPIDMINER: Rapid prototyping for complex data mining tasks. In Proceedings of the *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (KDD 2006).

Mislevy, R. J., & Haertel, G. D. (2006). Implications of evidence-centered design for educational testing. *Educational Measurement: Issues and Practice*, 25(4), 6–20.

Mislevy, R. J., & Riconscente, M. M. (2006). Evidence-centered assessment design. *Handbook of test development*, 61–90.

Murray, T. (1999). Authoring intelligent tutoring systems: An analysis of the state of the art. *International Journal of Artificial Intelligence in Education (IJAIED)*, *10*, 98–129.

Murray, T. (2003). EON: Authoring tools for content, Instructional strategy, student model and interface design. In *Authoring Tools for Advanced Technology Learning Environments* (pp. 309–339). Springer Netherlands.

Person, N. K., Graesser, A. C., Kreuz, R. J., & Pomeroy, V. (2003). Simulating human tutor dialog moves in AutoTutor. *International Journal of Artificial Intelligence in Education (IJAIED)*, 12, 23–39.

Rowe, J., Pokorny, B., Goldberg, B., Mott, B., & Lester, J. (2017). Toward Simulated Students for Reinforcement Learning-Driven Tutorial Planning in GIFT. In Proceedings of R. Sottilare (Ed.) 5th Annual GIFT Users Symposium. *US Army Research Laboratory*, Orlando, FL.

Salas, E. (2015). Team training essentials: A research-based guide. Routledge.

Schatz, S., Bowers, C., & Nicholson, D. (2009, October). Advanced situated tutors: Design, philosophy, and a review of existing systems. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 53, No. 26, pp. 1944–1948). Sage CA: Los Angeles, CA: SAGE Publications.

Simpson, E. (1972). The psychomotor domain. Washington DC: Gryphon House.

Sinatra, A. M. & Sottilare, R. (2016). Chapter 14 – Exploring the Diversity of Domain Modeling for Training and Educational Applications. In R. Sottilare, A. Graesser, X. Hu, A. Olney, B. Nye &

A. Sinatra (Eds.) *Design Recommendations for Intelligent Tutoring Systems: Volume 4 – Domain Modeling*. *Army Research Laboratory*, Orlando, Florida. ISBN: 978-0-9893923-9-6.

Soller, A. (2001). Supporting social interaction in an intelligent collaborative learning system. *International Journal of Artificial Intelligence in Education (IJAIED)*, *12*, 40–62.

Sottilare, R. (2012). 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.

Sottilare, R. A., Brawner, K. W., Goldberg, B. S. & Holden, H. K. (2012). The Generalized Intelligent Framework for Tutoring (GIFT). Concept paper released as part of GIFT software documentation. Orlando, FL: *U.S. Army Research Laboratory – Human Research & Engineering Directorate (ARL-HRED)*. Retrieved from: https://gifttutoring.org/attachments/152/GIFTDescription_0.pdf

Sottilare, R., Goldberg, B., Brawner, K., & Holden, H. (2012). A modular framework to support the authoring and assessment of adaptive computer-based tutoring systems (CBTS). In Proceedings of the *Interservice/Industry Training Simulation & Education Conference*, Orlando, Florida, December 2012.

Sottilare, R., Ragusa, C., Hoffman, M. & Goldberg, B. (2013). 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.

Sottilare, R., (2014). Using Learner Data to Influence Performance during Adaptive Tutoring Experiences. In Proceedings of the *8th International Conference on Augmented Cognition*. HCI International, Creta Maris, Heraklion, Crete, Greece, June 2014.

Sottilare, R. and LaViola, J. (2015). Extending Intelligent Tutoring Beyond the Desktop to the Psychomotor Domain: A survey of smart glass technologies. In Proceedings of the *Interservice/Industry Training Simulation & Education Conference*, Orlando, Florida, December 2015.

Sottilare, R., Burke, S., Johnston, J., Sinatra, A., Salas, E., and Holden, H. (2015). Antecedents of Adaptive Collaborative Learning Environments. In Proceedings of the *Interservice/Industry Training Simulation & Education Conference*, Orlando, Florida, December 2015.

Sottilare, R., Hackett, M., Pike, W., and LaViola, J. (2016). Adaptive Instruction for Medical Training in the Psychomotor Domain. In J. Cohn, D. Fitzhugh, and H. Freeman (Eds.) Special Issue: Modeling and Simulation Technologies to Enhance and Optimize the DoD's Medical Readiness and Response Capabilities of the *Journal for Defense Modeling & Simulation (JDMS).*

Sottilare, R., Brawner, K., Sinatra, A. & Johnston, J. (2017). An Updated Concept for a Generalized Intelligent Framework for Tutoring (GIFT). Orlando, FL: *US Army Research Laboratory*. May 2017. DOI: 10.13140/RG.2.2.12941.54244.

Sottilare, R. A., Burke, C. S., Salas, E., Sinatra, A. M., Johnston, J. H. & Gilbert, S. B. (2017). Towards a Design Process for Adaptive Instruction of Teams: A Meta-Analysis. *International Journal of Artificial Intelligence in Education.* DOI: 10.1007/s40593-017-0146-z.

Sottilare, R., Long, R., & Goldberg, B. (2017, April). Enhancing the Experience Application Program Interface (xAPI) to Improve Domain Competency Modeling for Adaptive Instruction. In Proceedings of the Learning @ Scale Conference, Cambridge, Massachusetts, April 20–21, 2017. DOI: http://dx.doi.org/10.1145/3051457.3054001.

Sottilare, R. A. & Ososky, S. (2017). Defining Complexity in the Authoring Process for Adaptive Instruction. In *Foundations of Augmented Cognition*. Springer International Publishing.

Sottilare, R. (2017a). Adaptive Instructional Methods to Accelerate Learning and Enhance Learning Capacity. *International Defense & Homeland Security Simulation Workshop of the I3M Conference*. Barcelona, Spain, September 2017.

Sottilare, R. (2017b). Challenges in Authoring, Instructional Management, and Evaluation Methods for Adaptive Instructional Systems. In *Proceedings of the American Education Research Association (AERA) Conference.* Technology, Instruction, Cognition & Learning Special Interest

Group (AERA TICL SIG) Symposium on Intelligent Tutoring Systems, BIG DATA–Learning Analytics, and Automated Humanlike Tutoring: Similarities and Differences in San Antonio, Texas, 28 April 2017.

Stevens, R. H., Galloway, T., Berka, C., & Sprang, M. (2009, July). Can neurophysiologic synchronies provide a platform for adapting team performance?. In International Conference on Foundations of Augmented Cognition (pp. 658–667). Springer Berlin Heidelberg.

Swartout, W. R., Gratch, J., Hill Jr, R. W., Hovy, E., Marsella, S., Rickel, J., & Traum, D. (2006). Toward virtual humans. *AI Magazine*, *27*(2), 96.

Traum, D. (2008). Talking to virtual humans: Dialogue models and methodologies for embodied conversational agents. *Modeling Communication with Robots and Virtual Humans*, 296–309.

VanLehn, K., Koedinger, K. R., Skogsholm, A., Nwaigwe, A., Hausmann, R. G., Weinstein, A., & Billings, B. (2007, July). What's in a step? Toward general, abstract representations of tutoring system log data. In *International Conference on User Modeling* (pp. 455–459). Springer Berlin Heidelberg.

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

Vargas González,, A., Williamson, B., LaViola, J., and Sottilare, R. (2017). Authoring Augmented Reality Scenarios for Intelligent Tutoring Systems with GIFT. In R. Sottilare (Ed., 2017) *5th Annual GIFT Users Symposium (GIFTSym5). Army Research Laboratory*, Orlando, Florida. ISBN: 978-0-9977257-1-1.

Woolf, B. P. (2010). A Roadmap for Educational Technology. *National Science Foundation # 0637190*

Zachary, W., Cannon-Bowers, J. A., Bilazarian, P., Krecker, D. K., Lardieri, P. J., & Burns, J. (1999). The Advanced Embedded Training System (AETS): An intelligent embedded tutoring system for tactical team training. *International Journal of Artificial Intelligence in Education* (IJAIED), 10, 257–277.

Zook, A., Riedl, M., Holden, H., Sottilare, R. & Brawner, K. (2012). Automated Scenario Generation: Toward Tailored and Optimized Military Training in Virtual Environments. In Proceedings of the Foundations of Digital Games Conference 2012, Raleigh, North Carolina, 30 May-1 June 2012.

ABOUT THE AUTHOR

Dr. Robert Sottilare leads adaptive instructional research within US Army Research Laboratory's Learning in Intelligent Tutoring Environments (LITE) Lab and is a co-creator of the Generalized Intelligent Framework for Tutoring (GIFT).

Robert A. Sottilare, Ph.D.

US Army Research Laboratory

12423 Research Parkway, Orlando, Florida 32826

e-mail: robert.a.sottilare.civ@mail.mil