ARL-SR-0336 ● SEP 2015

US Army Research Laboratory

Individual Learner and Team Modeling for Adaptive Training and Education in Support of the US Army Learning Model: Research Outline

by Greg Goodwin, Joan Johnston, Robert Sottilare, Keith Brawner, Anne Sinatra, and Arthur Graesser

NOTICES

Disclaimers

ARL-SR-0336 ● SEP 2015

US Army Research Laboratory

Individual Learner and Team Modeling for Adaptive Training and Education in Support of the US Army Learning Model: Research Outline

by Greg Goodwin, Joan Johnston, Robert Sottilare, Keith Brawner, Anne Sinatra
Human Research and Engineering Directorate, ARL

Arthur Graesser
University of Memphis Institute for Intelligent Systems, Memphis, TN

REPORT DOCUMENTATION PAGE					*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)* September 2015	**2. REPORT TYPE** Final
3. DATES COVERED (From - To)	September 2014–March 2015
4. TITLE AND SUBTITLE	Individual Learner and Team Modeling for Adaptive Training and Education in Support of the US Army Learning Model: Research Outline
5a. CONTRACT NUMBER	
5b. GRANT NUMBER	
5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)	Greg Goodwin, Joan Johnston, Robert Sottilare, Keith Brawner, Anne Sinatra, and Arthur Graesser
5d. PROJECT NUMBER	
5e. TASK NUMBER	
5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	US Army Research Laboratory ATTN: RDRL-HRT-T Aberdeen Proving Ground, MD 21005-5425
8. PERFORMING ORGANIZATION REPORT NUMBER	ARL-TR-0336
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	
10. SPONSOR/MONITOR'S ACRONYM(S)	
11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT	Approved for public release; distribution is unlimited.
13. SUPPLEMENTARY NOTES	

14. ABSTRACT

While human tutoring and mentoring are common teaching tools, current US Army standards for training and education are group instruction and classroom training, also known as “one-to-many” instruction. Recently, the US Army placed significant emphasis on self-regulated learning (SRL) methods to augment institutional training where Soldiers will be largely responsible for managing their own learning. In support of the US Army Learning Model and to provide affordable, tailored, SRL training and educational capabilities for the US Army, the US Army Research Laboratory is investigating and developing adaptive tools and methods to largely automate the authoring (creation), delivery of instruction, and evaluation of computer-regulated training and education capabilities. A major goal within this research program is to reduce the time and skill required to author, deliver, and evaluate adaptive technologies to make them usable by a larger segment of the training and educational community. This research includes 5 interdependent research vectors: individual learner and unit modeling, instructional management principles, domain modeling, authoring tools and methods, and evaluation tools and methods. This report (1 of 6 interdependent research outlines) focuses on individual learner and team modeling research for adaptive training and education. Learner and team models enable adaptive training systems to determine the most appropriate training content and/or methods for each individual or team. The research focuses on the structure of those models and how those models can be used to best adapt training. This outline describes past and current research on learner and team modeling and identifies research gaps that need to be addressed.

15. SUBJECT TERMS

GIFT, design, adaptive training, domain modeling, learner modeling, team modeling

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Robert Sottilare
Unclassified	Unclassified	Unclassified	UU	48	**19b. TELEPHONE NUMBER (Include area code)** 407-208-3007

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

Contents

List of Figures

List of Tables

Preface

This report is 1 of 6 interdependent research outlines in the Adaptive Training research program. Portions of this text, which originated in ARL-SR-0325,[1] appear in all 6 reports to ensure that readers get the same cross-cutting information.

[1] Sottilare R, Sinatra A, Boyce M, Graesser A. Domain modeling for adaptive training and education in support of the US army learning model—research outline. Aberdeen Proving Ground (MD): Army Research Laboratory (US); June 2015. Report No.: ARL-SR-0325.

1. Introduction

Training and education tools and methods must be of sufficient intelligence to understand the needs of individual learners and units of learners, to mitigate negative learner states, and to guide and tailor instruction in real-time to optimize learning. These tools and methods must also be affordable, effective, and easy to access and use. These requirements are enablers of the US Army Learning Model (ALM), which includes an emphasis on self-regulated learning (SRL) where Soldiers are expected to manage their own learning and career development through the growth of metacognitive (e.g., reflection), self-assessment, and motivational skills (Butler and Winne 1995). While SRL skills are difficult to train and develop, support may be provided to the learner through "adaptive training technologies" (tools and methods), which may be focused to guide learning and reinforce SRL principles.

To support ALM, the US Army Research Laboratory (ARL) has developed a program of research called "adaptive training", which includes 6 interdependent research areas or vectors: individual learner and unit modeling, instructional management principles, domain modeling, authoring tools and methods, evaluation tools and methods, and architectural and ontological support for adaptive training. The reports documenting these vectors expand the scope of the adaptive tutoring research described in ARL-SR-0284 (Sottilare 2013) to support ALM requirements in the mid-term and long-term evolution of training and educational technology: the Synthetic Training Environment and the Future Holistic Training Environment for Live and Synthetic.

This report (1 of 6 interdependent research outlines) focuses on Learner and Team Modeling research for adaptive training and education. Currently, the majority of intelligent tutoring systems (ITSs), a form of adaptive training tool to support one-to-one computer-based instruction, support well-defined domains in mathematics, physics, and software programming. Since Soldiers operate in more complex, dynamic, and ill-defined domains, it is necessary to expand the scope of adaptive training tools and methods to support training and education in these militarily-relevant domains. Learner and team models enable adaptive training systems to determine the most appropriate training content and/or methods for each individual or team. Research focuses on the structure of those models and how those models can be used to best adapt training. This report describes past and current research on learner and team modeling and identifies research gaps that need to be addressed.

2. Research Goals and Objectives

The goal of the research described in this report is to model militarily relevant training domains to support individually tailored and intelligently guided training experiences as prescribed by the ALM (US Army Training and Doctrine Command 2011). To provide guidelines, best practices, tools, models, and methods in support of this research goal, the following are the primary objectives of this research:

- Develop and evaluate competency models within the Generalized Intelligent Framework for Tutoring (GIFT) (Sottilare et al. 2012a). This includes developing a structure for a competency model, identifying variables or learner assessments that should be included within the competency model, and then evaluating which assessments have the biggest impact on training effectiveness.
- Examine ways of using a competency model to adapt training in terms of the content (Domain module) to be delivered and the training method (Pedagogy module).
- Determine the important variables that drive small-unit team performance and develop ways to measure and model those factors in an adaptive training system. This will entail both basic research on team processes and research to develop technologies to reliably and passively record these variables from team members.
- Determine how small-unit teams mature and improve their skills over time to determine possible adaptive tutoring strategies for helping teams to improve. Implement and evaluate using those strategies to develop small-unit teams in an adaptive tutoring environment.

This report examines the background and requirements for adaptive training capabilities along with research challenges, dimensions of learner and team modeling, desired end states, and interdependencies with other adaptive training research vectors.

3. Background

While human tutoring and mentoring are common teaching tools, current US Army standards for training and education are group instruction and classroom training—also known as one-to-many instruction. Group instruction and classroom training have been generally focused on acquiring and applying knowledge in proxies for live-training environments (e.g., desktop simulations, virtual simulations, constructive simulations, and serious games).

Classroom training, especially for complex topics, is often taught as a series of lists that the instructor goes through in a linear fashion (Schneider et al. 2013). This approach puts a heavy burden on the learner to build mental models and make conceptual connections. Using this instructional methodology may lead to varying degrees of success due to individual differences in skills, traits, and/or preferences. More complex, ill-defined, or dynamic tasks may be difficult to instruct in a classroom environment especially if the cognitive elements of the task require spatial interaction to develop/maintain skills (e.g., marksmanship). This is an area in which adaptive training systems may be of particular benefit. An adaptive training system can provide individually tailored instruction to many students at once, either in support of a live instructor or as an alternative to a live instructor. To provide this individually tailored instruction, the adaptive training system requires a learner model.

Small group instruction in live environments has also been used to assess application of knowledge and the development of skills. A standard feedback mechanism for US Army training is the after-action review (AAR), where significant decision points and actions are captured for a small group discussion that is conducted after the completion of a training event to help capture teachable moments and to aid Soldiers in reflecting on their recent training experiences.

Both classroom training and small-group instruction are manpower intensive; requiring teachers, mentors, and support staff to guide the Soldier's experience. Today, ITSs primarily guide learner training and education for cognitive tasks in well-defined domains (e.g., problem solving and decision-making tasks in mathematics and physics). Soldiers tend to perform cognitive, affective, psychomotor, and social tasks in both well-defined (e.g., building clearing) and ill-defined domains (e.g., leadership, resource allocation). ITSs generally provide static training (e.g., sitting at a desktop computer to train on a serious game) that falls short in matching the dynamic nature of many US Army operational tasks (e.g., psychomotor tasks),thereby reducing opportunities to develop and transfer skills to the operational environment.

Research is needed to understand the characteristics, similarities, and differences of Army training domains (cognitive, affective, psychomotor, social, and hybrid) to develop efficient and effective tools and methods to support self-regulated learning in complex, ill-defined, and physically dynamic military domains. This research is necessary to support development of many components of an adaptive training system to include the learner, domain, and pedagogical models needed to deliver this training via an ITS.

3.1 Self-Regulated Learning and the US Army Learning Model

In 2011, the US Army placed significant emphasis on the development of SRL skills with the expectation that new methods of instruction (e.g., ITSs) would augment institutional training (i.e., classroom and small group instruction). One-to-one human tutoring has been shown to be significantly more effective than one-to-many instructional methods (e.g., traditional classroom instruction [Bloom 1984; VanLehn 2011]). However, it is not practical nor is it affordable to have one expert human tutor to mentor each Soldier in the US Army for every required operational task. This alone signals the need for capabilities to support one-to-one, tailored training, and educational experiences.

Additionally, under the ALM, Soldiers are largely responsible for managing their own learning, but SRL skills are difficult to train and develop (Butler and Winne 1995; Azevedo et al. 2009; Graesser and McNamara 2010). We anticipate adaptive training tools and methods will fill this gap and will provide personalized guidance to acquire, apply, retain, and transfer knowledge and skills to the operational environment. This signals the need for a computer-regulated learning strategy to augment missing SRL skills; however, for this strategy to be realized adaptive training technologies must first become affordable, sufficiently adaptive, and easy to use.

3.2 Motivation for Research

A promising alternative to one-to-one human tutoring is one-to-one adaptive training tools that include ITSs. Meta-analyses and reviews support the claim that ITS technologies routinely improve learning over classroom teaching, reading texts, and/or other traditional learning methods. These meta-analyses normally report effect sizes (sigma [σ]), which refers to the difference between the ITS condition and a control condition in standard deviation units. The reported meta-analyses show positive effect sizes that vary from $\sigma = 0.05$ (Dynarsky et al. 2007) to $\sigma = 1.08$ (Dodds and Fletcher 2004), but most hover between $\sigma = 0.40$ and $\sigma = 0.80$ (Ma et al. in press; Fletcher 2003; VanLehn 2011; Graesser et al. 2012; Steenbergen-Hu and Cooper 2013, 2014). Our current best meta-meta estimate from all of these meta-analyses is $\sigma = 0.60$. This performance is comparable to human tutoring, which varies from between $\sigma = 0.20$ and $\sigma = 1.00$ (Cohen et al. 1982; Graesser et al. 2012), depending on the expertise of the tutor. Human tutors have not varied greatly from ITSs in direct comparisons between ITS and trained human tutors (VanLehn et al. 2007; VanLehn 2011; Olney et al. 2012).

Graesser et al. (in press) are convinced that some subject matters will show higher effect sizes than others when comparing any intervention (e.g., computer trainers, human tutors, group learning) to a control. It is difficult to obtain high-effect sizes for literacy and numeracy because these skills are ubiquitous in everyday life and habits are automatized. For example, Ritter et al. (2007) reported that the Cognitive Tutor for mathematics has shown an effect size of $\sigma = 0.30$–0.40 in environments with minimal control over instructors. Human interventions to improve basic reading skills typically report an effect size of $\sigma = 0.20$. In contrast, when the student starts essentially from ground zero, such as many subject matters in science and technology, then effect sizes are expected to be more robust. ITSs show effect sizes of $\sigma = 0.60$–2.00 in the subject matters of physics (VanLehn et al. 2005; VanLehn 2011), computer literacy (Graesser et al. 2004; Graesser et al. 2012), biology (Olney et al. 2012), and scientific reasoning (Millis et al. 2011; Halpern et al. 2012). As a notable example, the Digital Tutor (Fletcher and Morrison 2012) improves information technology by an effect size as high as $\sigma = 3.70$ for knowledge and $\sigma = 1.10$ for skills. The effect size attributed to improved instruction and improved domain knowledge have not been separated in this analysis. Such large effect sizes would never be expected in basic literacy and numeracy.

Overall, these are promising results and equate to an increase of about a letter grade improvement over traditional classroom instruction. While ITSs are a promising technology to support adaptive training for individuals in well-defined domains like mathematics, physics, and computer programming, the US Army requires the ability to develop and exercise Soldier skills in more ill-defined domains (e.g., leadership) and at the unit level (e.g., collaborative learning and team training). Developing and maintaining the ability to make effective decisions under stress and in complex environments is also desirable.

Adaptive systems by their nature require additional content and complexity to support tailored learning for each user and as a consequence have a very high development cost, a major barrier to adoption by the US Army. Adaptive systems are also insufficiently adaptive to support tailored, self-regulated training and educational experiences across a broad spectrum of military tasks as required by the ALM. Today, few ITS authoring tools are generalized across all of the domains requiring training, and no evaluation criteria or standards have been developed to promote reuse and interoperability among ITSs (Sottilare et al. 2012b). In other words, current adaptive systems are not yet intelligent enough to support the tailored instruction required by the US Army in the breadth of domains being trained, but there is a stable foundation of 50 years of science on which to grow an adaptive training and education capability for the US Army.

3.3 Adaptive Training and Education Definitions

To support the ALM and affordable adaptive training and educational capabilities for the US Army, ARL is investigating and developing adaptive tools and methods. A desired end-state is the automation of authoring (creation) processes, instruction, and evaluation of computer-regulated training and education capabilities to help build SRL skills and support mixed-initiative interaction. A major goal within this research program is to reduce the time–cost and knowledge–skill required to author, deliver, and evaluate adaptive technologies to make them usable by a larger segment of the US Army training and educational community.

Adaptive training and education research includes elements of adaptive tutoring, distributed learning, virtual humans, and training effectiveness evaluation. Sottilare (2013) provides additional detail on research specific to ITSs in ARL-SR-0284 (2013). The following definitions are provided for this section to distinguish between adaptive training and education elements and also to highlight their relationships:

Adaptive Tutoring: also known as intelligent tutoring; tailored instructional methods to provide one-to-one and one-to-many computer-guided experiences focused on optimizing learning, comprehension, performance, retention, reasoning, and transfer of knowledge and acquired skills to the operational environment.

Adaptive Tutoring Systems: also known as ITSs; the mechanism or technologies (tools and methods) to provide tailored training and educational experiences; adaptive tutoring systems respond to changing states in the learner and changing conditions in the training environment to optimize learning; adaptive tutoring systems anticipate and recognize teachable moments.

Virtual Humans: artificially intelligent visual representations of people that simulate or emulate cognitive, affective, physical, and social processes.

Distributed Learning: concurrent distribution of training and educational content to multiple users at the point-of-need in which content is intelligently selected to support learning, increased performance, and long-term competency in selected domains.

Training/Learning Effectiveness: evaluation of the impact of training and educational tools and methods on usability, learning, comprehension, performance, retention, reasoning, and transfer of knowledge and acquired skills to the operational environment.

Adaptive Training and Education Systems: a convergence of ITSs and external training and education capabilities (e.g., serious games, virtual humans, simulations) to support engaging experiences with reduced need for authoring (Sottilare 2015).

GIFT (Sottilare et al. 2012a): an open-source, modular architecture whose goals are to reduce the cost and skill required for authoring adaptive training and educational systems, to automate instructional delivery and management, and to develop and standardize tools for the evaluation of adaptive training and educational technologies.

Adaptive training and education research at ARL is being conducted across 6 interdependent research vectors: individual learner and unit modeling; instructional management principles; domain modeling, authoring tools and methods; evaluation tools and methods; and architectural and ontological support. This report (1 of 6 interdependent research outlines) focuses on learner and team modeling research for adaptive training systems with the goal of guiding learning in militarily relevant training and educational domains.

Soldiers operate in a variety of complex, dynamic, ill-defined domains where their ability to persevere in the face of adversity, adapt to their situation, collaborate, and think critically are key to the successful completion of their assigned missions. To develop and exercise these skills, it is paramount for Soldiers to train in challenging environments. Presently these few challenging training environments have been largely provided through manpower-intensive methods or systems with little ability to adapt instruction to support their learning needs. To illustrate this point, Franke (2011) asserts that through the use of case study examples, instruction can provide the pedagogical foundation for decision-making under uncertainty. However, this approach is limited in implementation by the expanse of potential cases that would need to be consistently updated and maintained to support large populations like the US Army.

Adaptive systems like ITSs have been shown to be effective in promoting learning in primarily static (e.g., learners seated at desktop computers) instructional settings within relatively simple, well-defined domains (e.g., mathematics, physics) for individual learners. For our purposes, static instruction includes cognitive, affective, or social training tasks where a desktop computer delivers instruction and where the physical movement of the learner is limited to activities that can be conducted while seated. For example, static instruction can effectively support cognitive tasks involving decision-making and problem-solving but are less effective for training tasks involving motion and perception (e.g., land navigation and marksmanship). Ideally, we desire portable adaptive instructional capabilities to go with Soldiers to support training and education at their point-of-need across a wide spectrum of US Army operational tasks. Research is needed to develop tools and methods to support broader domain modeling, which is representative of the full spectrum of US Army operational tasks. Standards, interoperability, and automation (e.g., automated scenario generation) (Zook et al. 2012) will likely play a significant role in making adaptive training practical. In this way adaptive training technologies will have the greatest impact on organizational learning in the US Army.

4. US Army Requirements for Adaptive Training Systems and Learner/Team Modeling

The Army Science and Technology (S&T) community uses Warfighter Outcomes (WFOs) as the authoritative source for identifying Warfighter needs. WFOs are used to share research and future technology solutions. In the training and education (T&E) domain, the adaptive T&E research program is targeting 4 specific requirements to support the evolution of US Army training: adaptive training and education systems; big data; training at the point-of-need; and artificial intelligence (AI).

4.1 Adaptive Training and Education Systems and Learner/Team Modeling

The primary gap to be addressed under this US Army requirement is the lack of adaptive systems (e.g., intelligent tutors) to support individual and collective (team or unit) training. The US Army needs an adaptive training and education capability that is persistent and easy to use/access with minimal startup time. There are also requirements to automate an informal AAR (also known as a postexercise critique) to reduce the time and skills needed to produce the AAR and improve its focus and quality. Another line of thought notes that the AI in ITSs could be used to facilitate rapid mission planning and course-of-action analyses as a job aid in operational contexts.

The major connection between the adaptive training and education requirement and the Learner/Team Modeling research vector is the need to develop competency models that are relevant to military training audiences, which will be used to drive tailored individual and team instruction. Developing such competency models will enable training delivery and management systems to track competency acquisition, sustainment, and decay on an individual and team level, making it possible to provide highly tailored training to optimize individual and unit readiness. Critical variables of interest in this research vector are individual and team learning, performance, retention, and transfer.

4.2 Big Data and Learner/Team Modeling

The primary gap to be addressed under this US Army requirement is that there is a lack of capability to handle and process large amounts of structured and unstructured data (also referred to as big data). One capability that is needed is a structured data analytics program linking individual data (e.g., achievements) to required long-term competencies in military occupational specialties (MOSs). This would allow Soldiers to understand where they rank in terms of experiences and achievements among other Soldiers in their MOS. It would also allow the US Army to identify specific experiences among successful Soldiers in that MOS and provide a model for other Soldiers in that MOS to follow. The data could also be used by course managers and instructors to continuously improve instruction and the mental models of both human and computer-based instructors. Finally, data collected on trainee learning and performance during adaptive training experiences could be used to facilitate Unit Training Management, where unit commanders would have access to empirical data to support unit training decisions.

The major connection between the Army's big data requirement and Learner/ Team Modeling is the ability to collect learner data, learner states, and training environment data to automatically update and maintain individual and unit-level models of competency. Big data will also allow Army course managers to identify best practices over time and to promote agile configuration management of instructional content, and effective strategies, tactics, and techniques.

4.3 Training at the Point-of-Need and Learner/Team Modeling

The primary gap to be addressed under this US Army requirement is the lack of an easily accessible, persistent, cost-effective, and low-overhead training environment. A capability is needed to bring training to Soldiers instead of Soldiers going to fixed training locations. This point-of-need training capability would be easily distributed, web-based, and built upon open-enterprise architecture in the cloud. US Army training and educational opportunities would be available on demand anywhere and anytime. However, it should be noted that the delivery mechanism (e.g., laptop computer, mobile device, and smart glasses) for adaptive training is critical in determining the limitations of the domain model scope and complexity. For example, it may be extremely difficult to train all the complexities of a psychomotor task in a desktop computer setting.

The major connection between point-of-need training and Learner/Team Modeling is the practicality of extending adaptive training beyond desktop or classroom applications. As technologies that can be used to deliver training become increasingly portable, the desire for point-of-need training will only increase. Individual and team models will need to be equally accessible and interoperable if that training is to be delivered in a manner that is tailored to the needs of those individuals and teams and can be supported by the technology delivering the instruction. This will be critical in acquiring information about individual learners and teams to support real-time decision making during instruction.

4.4 AI Capabilities and Learner/Team Modeling

The primary gap to be addressed under this US Army requirement is that the US Army lacks an automated capability to replicate the complexity and uncertainty of the operational environment. This gap specifically points to the lack of adaptiveness in virtual humans, intelligent tutoring systems, and other training capabilities. This gap leads to Soldiers developing training-response strategies that result in less challenging training over time along with lower engagement and lower levels of learning and transfer of skills to more challenging operational environments.

The major connection between AI capabilities and domain modeling involve the discovery and innovation of techniques to support a concept called, “automated scenario evolution”. AI capabilities are needed to support automated scenario evolution where AI drives the generation of new “child” scenarios from a single-parent scenario based on dimensions of that scenario and the state of the trainee. In this way, the authoring burden for highly complex training and educational domains may be reduced.

For example, consider a single scenario where dimensions include variable challenge levels based on 3 threats (i.e., low, moderate, high), 3 types of field-of-view (i.e., narrow, moderate, and wide) and clear line-of-sight (i.e., near, moderate, and far). AI could spawn 27 new child scenarios based on combinations of these variables. For this automated scenario evolution to work, the AI must have a clear understanding of the impact of any possible next scenario on the current learner or team state so that it can select the scenario that is most likely to improve the individual or team competency in the desired direction.

AI-based capabilities in adaptive training and education systems may also support data acquisition (sensing), natural language, problem-solving strategies, and perceptual/interaction mechanisms in the tutor.

5. Understanding the Dimensions of Learner/Team Modeling

There are 4 typical elements that compose ITSs, a prime example of an adaptive training and education system: a learner or trainee model, an instructional or pedagogical model, a domain model, and some type of user interface. The domain model typically includes an expert or ideal student model by which the adaptive system measures/compares/contrasts the progress of the learner toward learning objectives. The domain model also includes the training environment, the training task, and all of the associated instructional actions (e.g., feedback, questions, hints, pumps, and prompts), which could possibly be delivered by the adaptive system for that particular training domain. Typical interaction between the learner, the training environment, and the adaptive system (tutoring agents) is shown in Fig. 1.

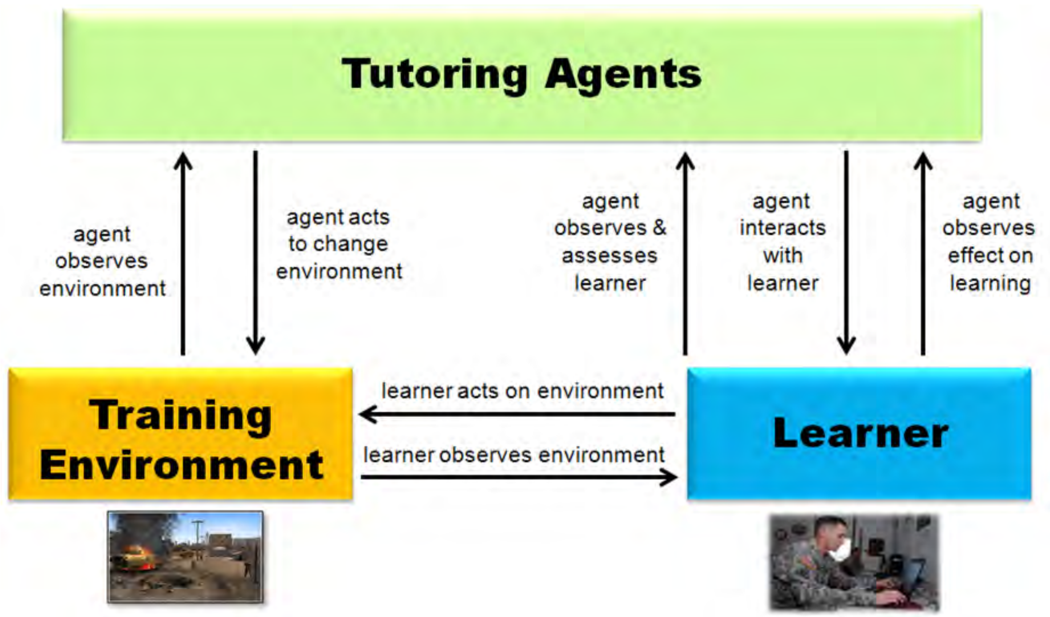

Fig. 1 Adaptive training interaction

Typical training systems examine the interaction between the learner and the training environment to measure progress toward learning objectives. The learner acts on the environment (e.g., opens a door or makes a choice to move into the room or stay outside) and then observes any changes or reactions within the environment. Adaptive systems add a layer of software-based tutoring agents that are designed to guide the learner in much the same way as a human tutor interacts with a learner. The tutoring agents observe the behaviors of the learner to assess their states (e.g., performance and attitudes) and interact with the learner to provide support, direction, and instruction. In addition, they track the effect of interactions on learning. Tutoring agents also interact with the training environment and may manipulate the environment to present more challenging or less challenging scenarios in response to the assessed state of the learner.

In order for the ITS to tailor training experiences, the system must have some “understanding” of the learner or team (i.e., a learner and/or team model). The ITS uses and updates these models as learners progress through the material. For example, if they master some concept or task, the models must be updated to reflect this. On the other hand, if the learners have difficulty with a concept or task the ITS needs to be able to understand why.

Understanding why the learner or team might have had difficulty is no simple task as the list of reasons could be quite extensive. Perhaps learners lose focus during the presentation of a key piece of information, lack some key prerequisite knowledge, or have a low aptitude for the topic or task. The list could go on and on.

All of these possible explanations require assessments of the learners. As can be seen from the earlier example, these assessments would need to include information about the learners’ backgrounds and experiences and measures of the learners during the training session. These assessments make up the learner and team models. In the following section, we discuss frameworks for describing these measures, first for the individual learner model and then for the team model.

5.1 An Assessment Framework for the Individual Learner Model

To begin, we will discuss a framework for individual learner assessments that might be found in a learner model. Such a framework is useful for understanding the ways in which learner assessments can be used. Knowing in advance how certain kinds of assessments can be used is of use in a generalized tutoring framework that needs to support a wide range of possible assessments.

In this report, we distinguish between the terms “measurement” and “assessment”. Measurement refers to behavioral or physiological data collected directly from the learner. Assessment is an inference about learner state or process (e.g., affective state, knowledge state). There are many ways to assess learner states and processes. For example, physiological data and facial expression data may be used to assess whether a learner is bored or frustrated. Alternatively, the ITS might simply ask the learner to self-report mood.

The framework in Section 5.2 describes learner assessments. Therefore, it does not include the measures that would be used to make those assessments. This is not to say that all measurement methods or technologies are equal, it is just that GIFT should remain agnostic to how assessments are made and should focus instead on how to use assessments. In this way, GIFT can incorporate new measurement methods or technologies without having to change the core learner or pedagogical models.

5.2 Types of Measures in the Assessment Framework

To better understand the range of the constructs that can be assessed, the framework shown in Fig. 2 serves as a good organizing rubric. In this figure, assessments are divided into 2 broad classes (pretraining assessments and in-training assessments) and 2 categories (content-dependent and content-independent).

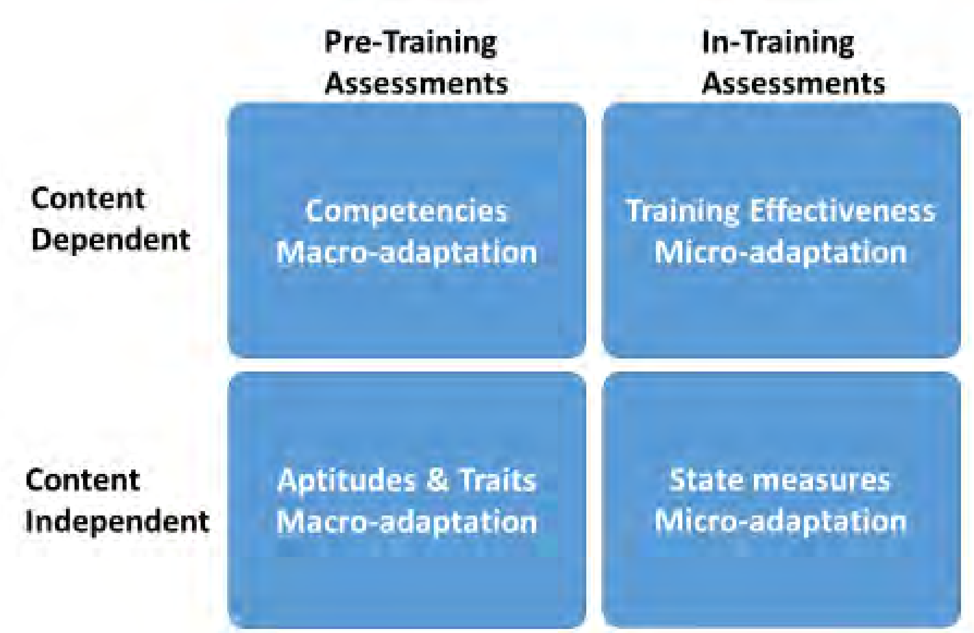

Fig. 2 Assessment framework for a learner model

There are 2 classes of assessments that ITSs are concerned with:

Pretraining Assessment: Measured before learning with the ITS begins. Some of these assessments are "traits" and may not change as a result of the training. Other assessments are competencies that may be updated by the training. This is what the learner brings to the training. These assessments are much more relevant to macro-adaptive training delivery. Many of these assessments could be derived from personnel records, learning and training management systems, etc. If they are not available from another source the ITS could administer the assessment immediately prior to the training session in the form of surveys and questionnaires.

In-Training Assessment: Assessed during training. These assessments are often "states" and are expected to change throughout training. This is what the learner takes away from the training. These assessments are collected by the training system and are relevant to micro-adaptive training delivery.

Across these 2 classes, assessments fall into 2 broad categories:

Content Dependent: As a pre-assessment, these are assessments of prior experience and skill or learner competencies that are relevant to the topic/skill being trained. At run-time, they are assessments of comprehension or skill execution directly associated with the training content (i.e., training effectiveness). Competencies could be updated as a result of training. In this way, the measures of training effectiveness would ultimately update the learner's overall competency.

Content Independent: As a pre-assessment, these are constructs like aptitude, personality traits, or physical abilities that impact a broad array of training topics/skills. As content-independent traits, it is unlikely they would be affected by the training, but they may impact macro-adaptive training delivery. At run-time, these are assessments that reflect learner states (e.g., fatigue, alertness, mood) that are not directly related to the topic/skill but may nevertheless affect the learner's current ability to learn.

It is worth noting that this framework incorporates the 3 types of measures outlined by Paneva (2006): transient states (in-training assessments), cumulative states (competencies), and enduring traits (aptitudes and traits). The difference is that the framework proposed earlier recognizes the content-dependent and content-independent nature of some of the measures and as a result, subdivides Paneva's transient measures into 2 categories.

Using this assessment framework has a couple of benefits. First of all, by understanding that there are different uses for each type of assessment, it is possible to think about ways that those uses might be standardized in GIFT modules. This might be especially true for content-independent measures. Second, it is useful in identifying research and technical challenges that affect certain types of assessments.

For example, in-training assessments are challenging because they must be frequently and rapidly assessed in a nonobtrusive way by the training system. Such assessments rely on measurement technologies like eye-trackers and physiological measures that can be expensive and may only be available in certain training facilities. Research and development will eventually bring the cost of these capabilities down and will likely increase their availability, but in the near term, this remains a problem.

In-training assessments of comprehension or skill are very domain specific and are expensive to develop and validate. Research will eventually reduce the time and cost of developing and validating these kinds of measures.

Finally, pretraining assessments can be challenging depending on the availability and currency of databases (e.g., personnel records, learner records) containing them. For example, access may be restricted if databases contain sensitive information like personally identifiable information, physical, and mental health data. Some databases may lack services to support interoperability (e.g., search, retrieve, replace functions). Finally, the records in a database may be out of date, incomplete, or lack the appropriate level of detail for a learner model in an ITS.

5.3 Domains of Learning

Domains of learning provide a very useful way to further delineate categories of assessments within the framework presented in Fig. 2. Bloom (1956) developed a well-known framework for describing learning domains: cognitive, psychomotor, and affective. Bloom subdivided the cognitive domain into 3 categories: factual, problem solving, and procedural. Others have updated and revised Bloom's categories, most significantly adding metacognitive to this list (Anderson and Krathwohl 2001).

Gagné (1970) proposed a similar set of domains, effectively breaking cognitive domain into 3 components (verbal, cognitive, and intellectual). His complete set included motor skills, verbal information (declarative knowledge), cognitive skills (decision making, problem solving, abstract reasoning, etc.), intellectual strategies (metacognitive skills), and attitudes.

Most disagreement deals with the types of categories within the cognitive domain. These distinction among the proposed cognitive categories are often philosophical in nature. That is, it is doubtful that experimentation will ever support one set of labels over another.

In the aggregate, there does seem to be agreement that cognitive, psychomotor, and affective are separate domains. Within the cognitive domain, common elements include declarative or factual knowledge, procedural knowledge, and something like metacognitive skills. Still others would argue that there is a fourth, social domain (e.g., Soller 2001). This could be added to the framework, but for simplicity we will limit our discussion to the cognitive, psychomotor, and affective learning domains in this report.

A competency can have components of all 3 domains. For example, a combat medic needs to understand human physiology and anatomy (cognitive domain), be able to apply first aid (psychomotor), and finally be able to remain calm and collected under combat conditions (affective domain).

5.4 Domains of Learning in the Assessment Framework

This report examines the specific categories of assessments from these domains of learning within our assessment framework. Table 1 displays examples of assessments for these 3 domains that could be included within each of the 4 quadrants of the assessment framework.

Table 1 Assessment framework with learning domains

Assessment Category	Domain	Pretraining Assessments (Macro-adaptation)	In-Training Assessments (Micro-adptation)
Content Dependent	Cognitive	Relevant prior cognitive experience/knowledge/training	Comprehension of concepts presented in the training
	Psychomotor	Relevant prior psyhomotor experience or training,	Measures of skill improvement
	Affective	Fears, likes, goals, attitudes relevant to the content.	Arousal and emotions in response to the training
Content Dependent	Cognitive	Intellect/aptitude, memory, Meta-cognitive skills	Alertness, cognitive workload
	Psychomotor	Physical strength, stamina, sensory acuity	Endurance, fatigue,
	Affective	Big 5 Personality traits	Arousal, emotions resulting from factors independent of training content

This provides a fairly comprehensive framework to organize the kinds of assessments that might be useful within an individual learner model. The 4 quadrants described in Fig. 2 still apply to Table 1. One of those quadrants, competencies, deserves further discussion.

5.5 Levels of Performance and Learner Competencies

Competence is the ability to do a job well. Competency is the set of knowledge, skills, and abilities that comprise competence in a specific job or role. Organizations define competencies in different ways. For example, the Army identifies twelve 21st century competencies and attributes. It further breaks these down into various learning outcomes, course outcomes, and learning objectives (US Army Training and Doctrine Command 2011). For the purposes of this report, we define a competency as specialized, job-specific skills, knowledge, and abilities that are developed over time and can be a result of both institutional and on-the-job training. Competencies are a very important element of the learner model.

Recall from the previously mentioned framework that competencies are comprised of pretraining and content-specific assessments. Competencies develop over time and often need continued practice to maintain. GIFT does not currently have a means of measuring competencies or of delivering training based on a learner's current competency level. In this section we discuss the structure of a competency framework that could be used within GIFT.

5.5.1 Cognitive Domain Levels

In the cognitive domain, the most recognizable model is Bloom's (1956) 6-stage model also known as Bloom's Taxonomy (knowledge, understanding, application, analysis, synthesis, and evaluation). These levels were later revised to remembering, understanding, applying, analyzing, evaluating, and creating (Anderson and Krathwohl 2001).

Even Merrill's First Principles theory (activate, demonstrate, apply, integrate) or his Component Display theory (rule, example, rule recall, practice) describe what could be called levels of performance.

Biggs and Collis (1982) developed the Structure of Observed Learning Outcomes. This consists of the following levels: unistructural (identify, name, simple procedure), multistructural (combine, describe, serial skills), relational (analyze, apply, relate, justify), and extended abstract (create, formulate, hypothesize).

5.5.2 Psychomotor Domain Levels

Pear (1927), a British psychologist, provided one of the earliest definitions of psychomotor skills. He felt that these skills had several features that distinguished them from aptitudes or habits. First, he said that a skill must be learned. Thus, walking on a tightrope would be a skill, whereas, walking on the ground would not. Additionally, he said that skills required an integration of many parts into a component whole. Thus, juggling several balls would be a skill, whereas tossing a single ball in the air would not. Finally, he said that psychomotor skills were primarily motor behaviors. This last requirement may be the most controversial because any skilled action relies on both cognitive and motor output. For example, a trained sniper must rely on practiced motor output when firing his or her rifle, but hitting the target also requires the sniper to correctly estimate distance to target and environmental factors like wind and temperature.

The development of psychomotor skills has produced several different models. As with the cognitive domain, there are common themes across the different models and the differences may be somewhat philosophical.

Simpson (1972) proposes 7 levels: perception (use sensory cues to guide motor behavior), set (readiness to act), guided response (early stage of learning a complex skill, imitation, trial and error), mechanism (intermediate stage of learning a complex skill, habitual response with some proficiency), complex overt (skillful performance, quick, accurate, coordinated), adaptation (skilled and able to adapt to special requirements), and origination (create new movement patterns to fit a situation).

Fitts and Posner (1967) propose a 3-stage model of expertise development. The first stage is the declarative stage. In this stage the learner has a factual understanding of the task and can describe steps involved, the purpose and end result of performing the task etc. Performance at this stage is highly variable and unskilled.

The second stage is the associative stage. In this stage the learner is consciously attempting to execute the task in accordance with their understanding of the task. The learner is developing "muscle memory" for the task. Performance at this stage is less variable and the learner's skill is improving. Cognitive load is high because the learner must concentrate on the execution of the task.

The third and final stage is the automaticity stage. In this stage the learner has developed muscle memory and can execute the task with little cognitive effort. Performance at this stage is stable (low variability) and skilled.

Finally, Dave (1970) proposes a 5-stage model: imitation (observing patterning behavior after someone else), manipulation (perform actions by memory or with instructions), precision (perform a skill with a high degree of precision), articulation (adapt a series of actions), naturalization (high level of performance, actions are automated).

5.5.3 Affective Domain Levels

The affective domain has to do with attitudes, beliefs, values, emotions, opinions, and motivation. Learning in the affective domain has to do with changing these affective components. Perhaps the best known description of the affective domain comes from Krathwohl, Bloom, and Masia (1964) who proposed the following 5 levels of the affective domain: receiving (awareness, willingness to hear), responds (react, comply), valuing (appreciates, behavior often impacted by values), organization (prioritize values, resolve conflicts between values), and finally internalized (consistent and pervasive impact of values).

5.6 Developing a Competency Model

As can be seen in the discussion of the prior section, there is little consensus about either the nature or number of abstract levels within learning domains. About the only thing that various theorists agree on is that there are levels. Although it may be difficult to define levels in the abstract, it is probably easier to define them for a specific competency. Indeed, the military frequently uses such graded metrics when assessing individuals and units. For example, a commonly used scale is trained, practice needed, and untrained . Another is novice, journeyman, and expert. Subject matter experts can provide specific indicators of individuals and units at these levels.

Putting aside the debate about the specifics of each level for the moment, we can propose a structure for a competency model as described in Table 2. For a given competency, a learner's progress can be modeled in terms of learning domain and level. For example, a medic may be at level 3 in the cognitive and psychomotor domains and level 2 in the affective domain. This information could then be used by an ITS to determine the most appropriate training and pedagogy for that learner.

Table 2 Competency model

Level	Domain		
	Psychomotor	**Affective**	**Cognitive**
Level 1	Measures/Training/ Pedagogy P1	Measures/Training/ Pedagogy A1	Measures/Training/ Pedagogy C1
Level 2	Measures/Training/ Pedagogy P2	Measures/Training/ Pedagogy A2	Measures/Training/ Pedagogy C2
Level 3	Measures/Training/ Pedagogy P3	Measures/Training/ Pedagogy A3	Measures/Training/ Pedagogy C3
Level n	Measures/Training/ Pedagogy Pn	Measures/Training/ Pedagogy An	Measures/Training/ Pedagogy Cn

Finally, there is no requirement that a given competency have training, measures, or pedagogies associated with all 3 domains of learning. Nor is there a requirement that there be an equal number of levels of each domain. There may be some competencies that have only 1 or 2 learning domains.

5.7 Standard Competencies

Eventually, it will be necessary to have common definitions of competencies and agreement on how they should be assessed, however this is not a basic research question. Most likely those definitions will need to come from the organization that is interested in tracking and training the competencies in its learner model.

For example, the Army has defined twelve 21st century competencies and attributes (US Army Training and Doctrine Command 2011). These competencies are further broken down in to general learning outcomes, branch-specific learning outcomes, course outcomes, and enabling or terminal learning objectives. Eventually, a learner model used for training Soldiers will need to align with this US Army Training and Doctrine Command-(TRADOC) defined competency hierarchy.

5.8 Areas of Research on Individual Learner Models for GIFT

The following are areas of research on individual learner models for GIFT that are currently being investigated:

Learner Affect: Current research is investigating the classification of within-environment affect detection (based on within simulation behaviors) and external acquisition of physiological measures. Automated classification is being compared with classification done by trained observers. The research focuses not only on whether such models can be produced but on how to do so in the most effective manner.

Metacognitive Skills: Current research is examining the feasibility of identifying and measuring metacognitive skills (e.g., reflection) utilized during instruction. The goal of this research is to identify the effective use of metacognitive skills and to develop methods of instruction to train metacognitive processes, which can be utilized across different domains. This efforts supports the US Army's desire to enhance the self-regulated learning habits of soldiers as part of the ALM.

Competency Modeling: Current research is examining methods to develop a competency map for various task domains (e.g., cognitive, affective, psychomotor). This element of a long-term learner model would be used to determine level of competency in training and educational domains based on past experiences (e.g., training). Performance measures will be used from multiple sources, not just GIFT-based tutoring. Performance data from new experiences will populate the long-term learner model and will be experience application programming interface (xAPI)-based, which has become a standard for Advanced Distributed Learning applications.

6. Individual Learner Model Research Challenges and Goals

It is expected that learners will continue to train in a range of live, virtual, and constructive environments. Furthermore, their skills, expertise, and competencies will be developed through both formal and on-the-job training. GIFT will therefore co-exist with all of these other training venues. For GIFT and other adaptive training systems to work efficiently in such an environment, it will be necessary for them to use and update interoperable learner databases. In this way, all adaptive training systems can readily access information about a learner's prior training and experiences and use that information to build learner models that describe learner competencies, experiences, traits, and aptitudes.

6.1 Goal 1 Develop and Evaluate a Competency Model within the GIFT Learner Module

This first goal is to implement a competency model within the GIFT architecture. Implementation of this model would have an impact on more than just the learner module. As noted in the previous discussion, the competency model would also impact the pedagogical module and the domain module and authoring tools.

For example, the domain module should identify the level to which the training should be associated and it should be identified as training to sustain a competency level, or training to increase the learner's skill to a higher competency level. In the same way, authoring tools for domain modules should enable the author to add such labels to the domain module.

The pedagogical module could map specific training approaches to competency domain levels. In fact, one of the main purposes for differentiating domains of learning and levels within them is to identify the most effective pedagogical approaches for the training (e.g., Gagné 1989). Thus, it would be possible for the pedagogical module to select a training method based on the specific domain and level of the competency being trained.

Three specific sub-goals for developing and evaluating competency models in GIFT are described as follows:

Develop a competency model for a specific competency that includes elements of all 3 domains of learning. Some work is currently planned to build competency models within GIFT using existing learner data; however, the focus is primarily on the cognitive learning domain. Research is needed to develop a competency model that includes psychomotor and affective domains as well cognitive domains. A good candidate for this work is marksmanship training. GIFT is currently being integrated into a marksmanship training simulator. A competency model for marksmanship training will enable GIFT to also encompass live marksmanship training.

Develop and evaluate ways to use competency model to recommend domain modules and pedagogy modules. Because the competency map can be used to identify appropriate training approaches and training content, it should be possible to use this map to select the appropriate domain and pedagogy modules in GIFT. This will require architectural changes to GIFT. Evaluation of this capability would entail research on the training effectiveness of this integration.

Implement skill retention models within the competency model**.** Competencies are developed over time, but without training, competency levels can decline. The competency model should be able to anticipate failures of memory based upon models of retention. In this fashion, mandatory training could be performed less often, and critical jobs skills can be refreshed as needed, leading to a better trained military.

6.2 Examine Ways to Develop Learner Models from Existing Data

There are an increasing number of repositories of learner data that could be leveraged to automatically build learner models. Referring to the framework described in Fig. 2, existing data would inform the pretraining assessment categories. Examples of data repositories include learning management systems, personnel management systems, and learner record stores (LRSs). The latter are based on the xAPI) standard developed by the Advanced Distributed Learning Co-Lab. These LRSs potentially contain very granular data describing multiple learner actions within a given training system. It should be possible to determine competency levels in cases where sufficient data exist.

A subgoal is to develop a standard for an interoperable learner model. While our focus is on developing the learner module within GIFT, as more training systems are developed with the ability to adapt to the needs of the learner, there will be a common need for learner models. It only makes sense to have all training systems update a single persistent, interoperable learner model, rather than requiring each system to “reinvent” a learner model for each student.

7. Team Behavior Modeling for Adaptive Tutoring

Soldiers and units operate in a variety of complex, dynamic, ill-defined domains where their ability to persevere in the face of adversity, adapt to their situation, collaborate, and think critically are key to the successful completion of their assigned missions. To develop and exercise these skills, it is paramount for Soldiers to train in challenging and effective learning environments. Currently, developing such experience and teamwork takes considerable resources including time and money, trips to firing ranges, and unit deployments to training centers. Presently, these few environments have been largely provided through manpower-intensive methods or systems with little ability to adapt to their learning needs.

TRADOC released the ALM to guide a revolution in Army training (US Army Training and Doctrine Command 2011). The ALM calls for a total integrated learning enterprise that provides the capabilities to ensure education and training is available and effective in developing the 21st century Soldier competencies that includes teamwork and collaboration because the squad will remain the foundation and cornerstone of the Army.

> With advances in technology and in warfighting, the squad leader is given access to more data and more firepower than ever before. Furthermore, tactical small unit leaders must have improved situational awareness, judgment, and emotional maturity to determine if, when, and how the application of lethal force would best support the mission. Leaders build teams, seek multiple perspectives, alternative viewpoints, and manage team conflict. Squad members must excel at teamwork. Effective team members understand team dynamics, and take appropriate action to foster trust, cohesion, communication, cooperation, effectiveness, and dependability within the team (p. 42).

The ALM recommends that S&T should define effective tactical small units, and determine how they are formed more quickly and efficiently, with higher levels of experience in fractions of the time.

One of the challenges of conducting team research is in differentiating between individual performance of team members in the context of a team and team performance. Team behaviors include communication, information exchange, supporting behaviors, trust, and leadership. Individual behaviors include the unique tasks that each member must perform.

For example, suppose one squad of Soldiers engaged the enemy by simply having all members open fire, aiming at whatever targets presented themselves. Now suppose another squad used one fire team to suppress the enemy (fire on their position to force them to take cover) while the other fire team maneuvered into a flanking position. In the first example, the squad was little more than a group of individuals executing individual skills (i.e., engaging targets with their weapons). There was no communication or coordination of their efforts. In the second example, the squad members coordinated their efforts to enable one fire team to gain an advantageous position. They still executed similar individual skills but they also utilized such team behaviors as leadership, initiative, communication, supporting behaviors, and trust.

Key team tasks to measure when examining team performance include the following: identify the objective, elaborate the objective, plan to accomplish the objective, and execute the plan. Key skills include those mentioned previously (i.e., information exchange, communication, supporting behaviors, initiative and leadership, and trust). The details of how one would measure these in the context of a team task depend on the details of the team task.

Detecting, measuring, and providing effective feedback for team behaviors is difficult even for skilled human trainers and coaches. Developing this set of capabilities in an intelligent tutoring system is even more challenging. While theory and empirical research on modeling aspects of team effectiveness and performance is substantial (Salas et al. 2015), such research involving Soldiers or squads is extremely scarce and limited in scope (Sottilare et al. 2011; Sottilare 2012). Similarly, theory and empirical research on military team training has matured for small units, but much of the focus has not been on Soldiers or squads (; Sottilare et al. 2011; Sottilare 2012). Adaptive team tutoring has almost exclusively focused on Navy combat teams but has not matured beyond applied research and a few advanced technology demonstrations (Zachary et al. 1998; Lyons and McDonald 2001; Dorsey et al. 2009; Rothrock et al. 2009;). For team tutoring to be effective it is important to understand the kinds of teams and team tasks that can benefit from intelligent team tutors.

We conducted a literature review and meta-analysis of the existing scientific literature to identify the factors that influence team outcomes and the statistical relationships of those factors. Findings from this enabled the development of an initial set of behavioral markers and metrics that could be embedded within an adaptive tutor simulation to accurately assess teamwork processes and performance throughout tutoring. Based on the analysis, an initial set of guidelines and recommendations for effective team training strategies were developed. The final stage of this effort is ongoing and will be a refined team design architecture that is built upon the resulting best practices and principles derived from the initial review.

Additionally, in 2014 we began designing an adaptive team tutoring research test bed that includes a team modeling architecture and behavioral markers. The first phase of the research involves adapting the VBS3 (Virtual Battlespace 3) Games for Training simulation to model a 2-person reconnaissance task and instrument it to collect individual actions that represent affect, trust, team work, and team taskwork behaviors. Experiments are planned to establish a baseline of team performance assessment and prepare the testbed for tutoring and feedback strategies. Over time we will examine increasingly complex task domains and numbers of team members to test the robustness of team modeling and the adaptive tutor capabilities.

8. Team Modeling Research Goals and Challenges

8.1 Determine the Important Variables and Metrics Needed for Modeling Small Unit Team Processes and Performance Outcomes That Can Be Used in Adaptive Tutoring

While initial team models have been theorized, work is needed to identify a complete design architecture, including behavioral markers and metrics that can be used to model team processes and performance in adaptive tutoring (Sottilare et al. 2011; Sottilare 2012). This design architecture must be based on the science behind teamwork and team performance.

- Research is needed to determine the most important team factors that need to be modeled and assessed for team tutoring.
- Research is needed to determine whether these factors can be independent of team type and task type so that team tutoring can optimize team development and maturation.

8.2 Design Simulation Technologies So They Can Accurately Capture, Assess, and Model Small Unit Team Behaviors for Adaptive Tutoring

Figure 3 is a notional model of an adaptive tutoring effect chain for team tutoring. It illustrates the need to design adaptive tutoring technologies to accommodate the complexity of data collection and analytics required for assessing team members to inform team states from moment-to-moment for an adaptive tutoring approach. Research is needed to accomplish the following:

- Develop low-cost passive sensing technologies of team member behaviors and internal state as it relates to the team state.
- Develop technologies that can classify the necessary team behaviors, including affect and trust (and other internal states), with sufficient accuracy to make predictions of team performance.

- Identify and implement effective instructional strategies, balancing a focus on individual instruction versus team instruction, including determining the feedback strategies needed (e.g., real-time or near real-time) to improve team performance.

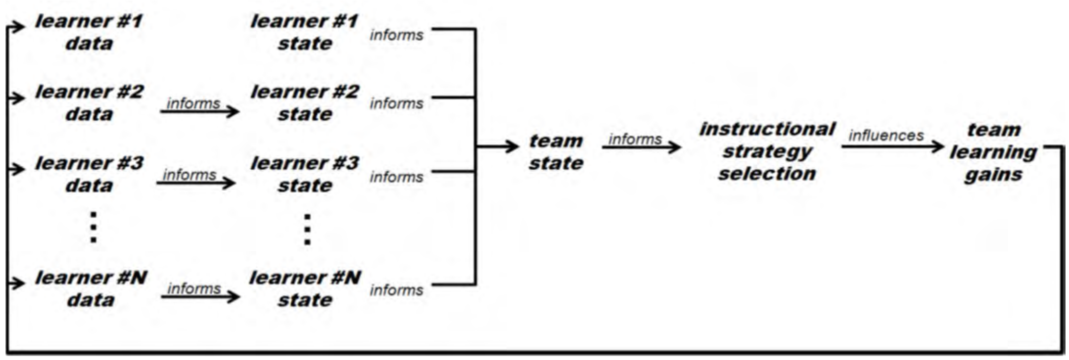

Fig. 3 Notional adaptive tutoring learning effect chain for team tutoring

8.3 Understand How Small Unit Teams Mature Over Time So That Modeling and Tutoring Can Support Those Changes

Past research has found that a combat team training curriculum should focus on developing team skills in a stepwise fashion so that a team of experts can become an expert team. Taskwork skills (combat mission execution and team decision making) should be trained first. Training for team tactical decision making should focus on the team decision making cycle; the 4 major team taskwork skill areas are identification, elaboration (critical thinking), planning, and execution (Johnston et al. 2013). Once these skills are developed, then teamwork skills (information exchange, initiative/leadership, backup and error correction, and communication structure) can be developed. The common mode of feedback for learning is through a team leader who is expert in leading a team self-correction AAR (Smith-Jentsch et al. 1998; Smith-Jentsch et al. 2008).

Furthermore, team researchers have proposed that teams develop over time through a maturation process (e.g., forming, storming, norming, etc.), and have hypothesized that taskwork and teamwork skills mature at different rates, and that team leaders must change their leadership strategies during this maturation process to enable teams to improve (Kozlowski et al. 2008).

- Research is needed to identify how teams develop through stages of maturation and how it should be measured.

- Research is needed to determine what team tutoring strategies might accelerate more rapid team maturation.

8.4 Develop Distributed Adaptive Tutors That Accurately Capture, Assess, and Model Small Unit Team Behaviors for Training in Collective Training Exercises

Squads consist of multiple teams (i.e., fire teams) and they in turn are embedded in multi-team systems (i.e., platoons, company, and higher). Team integration based on mission tasks happens within a squad, across squads within a platoon, and squads working across platoons. Unit training involves developing mission-task competence for these interactions (Fowlkes et al. 2005).

- Research is needed to understand and accurately model small unit team integration for collective missions.
- Research is needed to determine the best approach for adaptive tutoring for collective training.
- Research is needed to develop distributed adaptive tutors that accurately capture, assess, and model small unit team behaviors for training in collective training exercises.

9. Interdependencies with Other Adaptive Training Research Vectors

This section examines interdependencies between Learner and Team Modeling and the other 5 adaptive training research vectors (Fig. 4). This discussion forms the basis for the sequencing of research and ultimately bringing adaptive training capabilities into a state of practice.

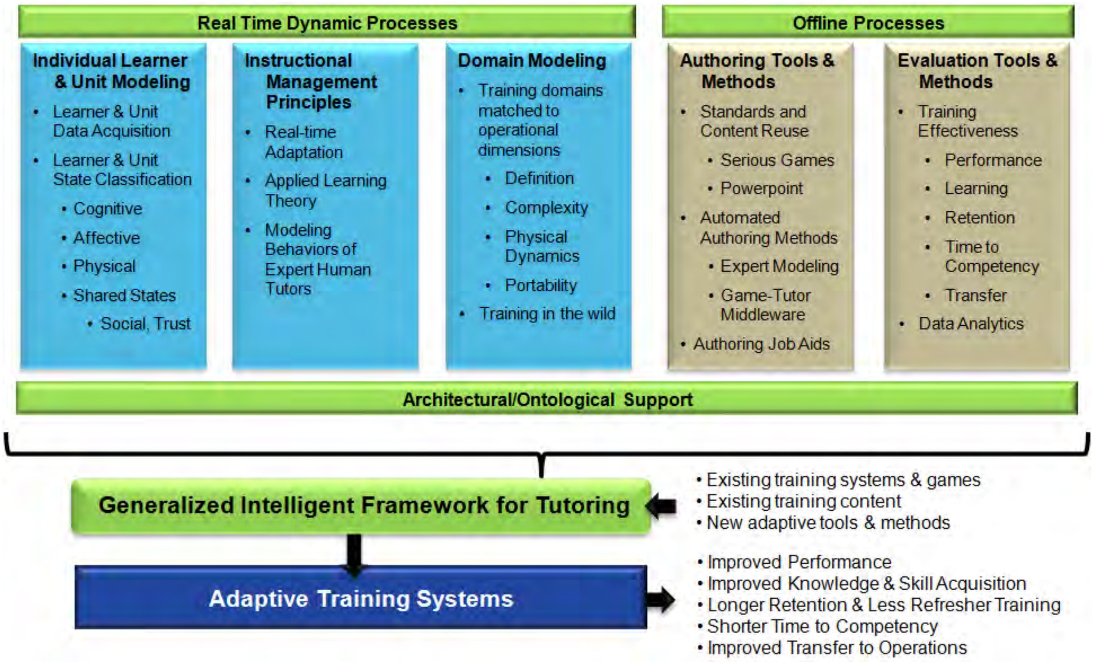

Fig. 4 Adaptive training research vectors

Accurate methods to classify individual and team learner states are a necessary precursor to selecting optimal instructional strategies as noted in the learning effect models for individual learners (Fig. 3) and teams of learners (Fig. 4). In turn instructional strategies along with instructional context are necessary precursors to selecting optimal instructional tactics and ultimately significant effect on desired outcomes: learning, performance, retention, and transfer.

9.1 Domain Modeling and Learner/Team Modeling

Adaptive training systems are learner-centric systems. Independent of the domain under training, accurate modeling of the learner is critical to driving instructional decisions in adaptive training systems, but collection and maintenance of this data may be costly so it is necessary to select measures and states which significantly impact our desired outcomes: learning, performance, retention, and transfer.

As competency models are implemented within GIFT, it will be necessary to associate the appropriate domain models with different levels of the competency learning domains. In this way the ITS can identify which content to deliver to the learner appropriate to his or her level of competency.

9.2 Automated Instruction and Learner/Team Modeling

In GIFT, instructional management takes place in 2 modules/processes within the learning effect model. One process is instructional strategy selection within the pedagogical module. The second is within the domain module where specific tactics or actions are selected based on the strategy selection and instructional context.

As with the domain module, pedagogical strategies should be associated with the components of the competency model. As noted earlier, one of the main reasons theorists originally described domains of learning and levels within those domains was to provide a framework for determining the most appropriate instructional methods for different learning outcomes (Gagné 1989).

9.3 Authoring Tools and Learner/Team Modeling

Authoring tools and methods will produce the content and tactics included in the domain module of GIFT. As such, authoring tools must be able to provide the labels or markers that associate the domain content and tactics with the relevant elements of the competency model. Authoring tools will also be needed to create custom competency models.

9.4 Evaluation and Learner/Team Modeling

Research is needed to determine the best ways of evaluating competency levels of learners and teams. This is essential to both determine the competency level of learners and to evaluate the ability of different training methods and content to sustain or improve learner competency.

9.5 Architecture and Learner/Team Modeling

The architecture supporting learner modeling needs to exist in 2 fashions: in data capture and storage, and in data communication and usage. Instructional strategies based upon learner modeling information (performance, traits, states, etc.) may exist only after the data are successfully captured and communicated to the instructional models responsible for these decisions. The development and tracking of this long-term model allows for the research of its application.

10. Conclusions

The GIFT Intelligent tutoring system has been a fruitful tool for research on adaptive training for the past 5 years. Moving forward, research on the learner and team models will be central to the further development of this system.

Accurate methods to classify individual and team learner states are a necessary precursor to selecting optimal instructional strategies for individuals and teams of learners. To the degree that such optimal strategies can be selected, GIFT will be able to maximize its impact on learning, performance, retention, and transfer

Research and development on competency modeling will enable GIFT to more easily adapt its training to individual learners and it will enable GIFT to operate seamlessly in an Army Live, Virtual, and Constructive training ecosystem. Improving the interactions between the learner module and pedagogical and authoring modules has the potential to facilitate authoring and improve selection of appropriate pedagogical approaches.

Team modeling remains a particularly challenging area of research. Key areas of future research should include developing measures for team processes and performance; designing technologies that assess and model small unit team behaviors; understand how small unit teams mature over time; and finally developing distributed tutors that can assess and train teams of teams. Although team tutoring research is still in its early stages, the importance of team tutoring research for the Army cannot be overstated as virtually every Army function involves some level of teamwork.

11. References

Anderson LW, Krathwohl DR, editors. A taxonomy for learning, teaching and assessing: a revision of bloom's taxonomy of educational objectives: complete edition. White Plains (NY): Longman; 2001.

Azevedo R, Witherspoon A, Graesser AC, McNamara DS, Chauncey A, Siler E, Cai Z, Lintean M. MetaTutor: analyzing self-regulated learning in a tutoring system for biology. In: Dimitrova V, Mizoguchi R, Du Boulay B, Graesser AC, editors. Artificial intelligence in education: building learning systems that care: from knowledge representation to affective modelling. Amsterdam (The Netherlands): IOS Press; 2009. p. 635–637.

Biggs JB, Collis KF. Evaluating the quality of learning – the SOLO taxonomy. New York (NY): Academic Press; 1982.

Bloom, BS. Taxonomy of educational objectives: handbook I: cognitive domain. White Plains (NY): Longman; 1956.

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

Butler DL, Winne PH. Feedback and self-regulated learning: a theoretical synthesis. Review of Educational Research. 1995;65:245–281.

Cohen PA, Kulik JA, Kulik CLC. Educational outcomes of tutoring: a meta-analysis of findings. American Educational Research Journal. 1982;19:237–248.

Dave RH. Psychomotor levels. 1970. In: Armstrong RJ, editor. Developing and writing behavioral objectives. Tuczon (AZ): Educational Innovators Press; 1970.

Dodds P, Fletcher JD. Opportunities for new "smart" learning environments enabled by next generation web capabilities. Journal of Educational Multimedia and Hypermedia. 2004;13:391–404.

Dorsey D, Russell S, Keil C, Campbell G, Van Buskirk W, Schuck P. Measuring teams in action: automated performance measurement and feedback in simulation-based training. In: Salas E, Goodwin GF, Burke CS, editors. Team effectiveness in complex organizations: cross-disciplinary perspectives and approaches. New York (NY): Routledge; 2009. p. 351–381.

Dynarsky M, Agodini R, Heaviside S, Novak T, Carey N, Camuzano L, Sussex W. Effectiveness of reading and mathematics software products: findings from the first student cohort; March Report to Congress 2007. [accessed 2015 May 15]. http://ies.ed.gov/ncee/pdf/20074005.pdf.

Fitts PM, Posner MI. Human performance. Belmont (CA): Brooks/Cole; 1967.

Fletcher JD. Evidence for learning from technology-assisted instruction. In: O'Neil HF, Perez R, editors. Technology applications in education: a learning view. Mahwah (NJ): Erlbaum; 2003. p. 79–99.

Fletcher JD, Morrison JE. DARPA digital tutor: assessment data. Alexandria (VA): Institute for Defense Analyses; (IDA Document D-4686). 2012. [accessed 2015 May]. http://www.acuitus.com/web/pdf/D4686-DF.pdf.

Fowlkes J, Owens J, Hughes C, Johnston JH, Stiso M, Hafich A, Bracken K. Constraint-directed performance measurement for large tactical teams. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting; 2005 Sep 26–30; Orlando, FL. Thousand Oaks (CA): Sage Publications: 2005;49(25):2125–2129.

Franke D. Decision-making under uncertainty: using case studies for teaching strategy in complex environments. Journal of Military and Strategic Studies. 2011;13(2): 1–21. [accessed 2015 May]. http://ww.jmss.org/jmss/index.php /jmss/article/viewFile/385/398.

Gagné RM. The conditions of learning. 2nd edition. New York (NY): Holt, Rinehart, and Winston; 1970.

Gagné RM. Studies of learning. Tallahassee (FL): Learning System Institute; 1989.

Graesser AC, Hu X, Nye B, Sottilare R. Intelligent tutoring systems, serious games, and the generalized intelligent framework for tutoring (GIFT). In: O'Neil HF, Baker EL, Perez RS, editors. Using games and simulation for teaching and assessment. Oxon (UK): Routledge; Abingdon. In Press.

Graesser AC, Lu S, Jackson GT, Mitchell, HH, Ventura M, Olney A, Louwerse MM. AutoTutor: A tutor with dialogue in natural language. Behavioral Research Methods, Instruments and Computers. 2004;36:180–193.

Graesser AC, McNamara DS. Self-Regulated learning in learning environments with pedagogical agents that interact in natural language. Educational Psychologist. 2010;45(4):234–244.

Graesser AC, Conley M, Olney A. Intelligent tutoring systems. In: Harris KR, Graham S, Urdan T, editors. APA Educational Psychology Handbook: vol. 3. Applications to Learning and Teaching. Washington (DC): American Psychological Association; 2012. p. 451–473.

Halpern DF, Millis K, Graesser AC, Butler H, Forsyth C, Cai Z. Operation ARA: A computerized learning game that teaches critical thinking and scientific reasoning. Thinking Skills and Creativity. 2012;7:93–100.

Johnston JH, Fiore SM, Paris C, Smith CAP. Application of cognitive load theory to develop a measure of team cognitive efficiency. Military Psychology. 2013;25(3):252.

Kozlowski SWJ, Watola DJ, Jensen JM, Kim BH, Botero IC. Developing adaptive teams: a theory of dynamic team leadership. In: Salas E, Goodwin GF, Burke CS, editors. Team effectiveness in complex organizations: cross-disciplinary perspectives and approaches. New York (NY): Routledge; 2008. p. 113–155.

Krathwohl DR, Bloom BS, Masia BB, editors. Taxonomy of educational objectives: handbook II: affective domain; New York (NY): David McKay Co.; 1964.

Lyons DM, McDonald DP. Advanced embedded training with real-time simulation for Navy surface combatant tactical teams. In: Smith M, Salvendy G, editors. Systems, social and internationalization design aspects of human computer interaction. Hillsdale (NJ): Lawrence and Erlbaum Associates; 2001.vol. 2.

Ma W, Adesope OO, Nesbit JC. Intelligent tutoring systems and learning outcomes: a meta-analytic survey. Journal of Educational Psychology. In Press.

Millis K, Forsyth C, Butler H, Wallace P, Graesser AC, Halpern D. Operation ARIES! A serious game for teaching scientific inquiry. In: Ma M, Oikonomou A, Lakhmi J, editors. Serious games and edu-tainment applications. London (UK): Springer-Verlag; 2011. p. 169–196.

Olney AM, Person NK, Graesser AC. Guru: designing a conversational expert intelligent tutoring system. In: Boonthum-Denecke C, McCarthy P, Lamkin T, editors. Cross-disciplinary advances in applied natural language processing: issues and approaches. Hershey (PA): Information Science Publishing; 2012. p. 156–171.

Paneva D. Use of ontology-based student model in semantic-oriented access to the knowledge in digital libraries. Presented at HUBUSKA Fourth Open Workshop Semantic Web and Knowledge Technologies Applications; 2006 Sep 12; Varna (Bulgaria). p. 31–41.

Pear TH. Skill. Personnel Journal. 1927;5:478–489.

Ritter S, Kulikowich J, Lei P, McGuire CL, Morgan P. What evidence matters? A randomized field trial of cognitive tutor algebra I. In: Hirashima T, Hoppe U, Young SS, editors. Supporting learning flow through integrative technologies. Amsterdam (The Netherlands): IOS Press; 2007. p. 12–20.

Rothrock L, Cohen A, Yin J, Thiruvengada H, Nahum-Shani I. Analyses of team performance in a dynamic task environment. Applied Ergonomics. 2009;40(4):699–706.

Salas E, Shuffler ML, Thaye, AL, Bedwell WL, Lazzara EH. Understanding and improving teamwork in organizations: a scientifically based practical guide. Human Resource Management. 2015;54(4):599–622.

Schneider B, Wallace J, Blikstein P, Pea R. Preparing for future learning with a tangible user interface: the case of neuroscience. Learning technologies, IEEE Transactions. 2013;6(2):117–129.

Simpson E. The classification of educational objectives in the psychomotor domain: the psychomotor domain. Washington (DC): Gryphon House; 1972. Vol. 3.

Smith-Jentsch KA, Cannon-Bowers JA, Tannenbaum SI, Salas E. Guided team self-correction: impacts on team mental models, behavior, and effectiveness. Small Group Research. 2008;39:303–327.

Smith-Jentsch KA, Zeisig RL, Acton B, McPherson JA. Team dimensional training. In: Cannon-Bowers JA, Salas E, editors. Making decisions under stress: implications for individual and team training. Washington (DC): American Psychological Association; 1998. p. 271–297.

Soller A. Supporting social interaction in an intelligent collaborative learning system. International Journal of Artificial Intelligence in Education. 2001;12(1):40–62.

Sottilare R. Considerations in the development of an ontology for a generalized intelligent framework for tutoring. International Defense and Homeland Security Simulation Workshop in Proceedings of the International Modeling and Simulation Multiconference; 2012 Sep 21; Vienna, Austria. Genoa (Italy): Department of Mechanical, Energy, Logistics Engineering and Engineering Management (DIME) University of Genoa; c2012.

Sottilare R. Special report: adaptive intelligent tutoring system (ITS) research in support of the army learning model - research outline. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2013. Report No.: ARL-SR-0284. Also available at http://www.arl.army.mil/www/default.cfm?technical_report=6942.

Sottilare R. Challenges in moving adaptive training and education from state-of-art to state-of-practice. In: Conati C, Hefferman N, Mitrovic A, Verdejo MF, editors. Proceedings of Developing a Generalized Intelligent Framework for Tutoring (GIFT): Informing Design through a Community of Practice Workshop at the 17th International Conference on Artificial Intelligence in Education (AIED 2015); 2015 Jun 22–26; Madrid, Spain. Heidelberg (Germany): Springer-Verlag; c2015.

Sottilare RA, Brawner KW, Goldberg BS, Holden HK. The generalized intelligent framework for tutoring (GIFT). Orlando (FL): Army Research Laboratory (US); Human Research and Engineering Directorate (HRED); 2012a [accessed 2015 May]. https://gifttutoring.org/attachments/152 /GIFTdescription_0.pdf.

Sottilare R, Goldberg BS, Brawner KW, Holden HK. 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; 2012 Dec 3–6; Orlando, FL. Arlington (VA): National Defense Industrial Association; c2012b.

Sottilare R, Holden H, Brawner K, Goldberg B. Challenges and emerging concepts in the development of adaptive, computer-based tutoring systems for team training. Proceedings of the Interservice/Industry Training Simulation & Education Conference; 2011 Nov 30–Dec 4; Orlando, FL. Arlington (VA): National Defense Industrial Association; c2011.

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. 2013;105(4):970–987.

Steenbergen-Hu S, Cooper H. A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. Journal of Educational Psychology. 2014;106:331–347.

US Army Training and Doctrine Command. The United States Army learning concept for 2015. Fort Monroe (VA): Headquarters, US Army Training and Doctrine Command; 2011 [accessed 2015 May]. http://www.tradoc .army.mil/tpubs/pams/tp525-8-2.pdf.

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

VanLehn K, Graesser AC, Jackson GT, Jordan P, Olney A, Rosé CP. When are tutorial dialogues more effective than reading? Cognitive Science. 2007;31(1):3–62.

VanLehn K, Lynch C, Schulze K, Shapiro JA, Shelby R, Taylor L, Weinstein A, Wintersgill M. The Andes physics tutoring system: lessons learned. International Journal of Artificial Intelligence and Education. 2005;15(3):147–204.

Zachary W, Cannon-Bowers JA, Burns J, Bilazarian P, Krecker D. An advanced embedded training system (AETS) for tactical team training. In: Goettl BP, Halff HM, Redfield CL, Shute VJ, editors. Paper presented at the Fourth International Conference on Intelligent Tutoring Systems; 1998 Aug 16–19; San Antonio, TX. Heidelberg (Germany); Springer-Verlag; c1998.

Zook A, Lee-Urban S, Riedl M, Holden H, Sottilare R, Brawner K. Automated scenario generation: toward tailored and optimized military training in virtual environments. Paper presented at the Foundations of Digital Games ’12; 2012 May 30–1 Jun; Raleigh, NC.

List of Symbols, Abbreviations, and Acronyms

AAR	after-action review
AI	artificial intelligence
ALM	Army Learning Model
ARL	US Army Research Laboratory
GIFT	Generalized Intelligent Framework for Tutoring
ITS	Intelligent Tutoring System
LRS	learner record stores
MOS	military occupational specialty
S&T	Science and Technology
SRL	self-regulated learning
T&E	training and education
TRADOC	US Army Training and Doctrine Command
WFO	Warfighter Outcome
xAPI	experience application programming interface

1 (PDF) DEFENSE TECHNICAL
INFORMATION CTR
DTIC OCA

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

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

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

1 (PDF) ARMY RSCH LAB – HRED
RDRL HRM DG K GUNN
BLDG 333
PICATINNY ARSENAL NJ 07806-5000

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

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

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

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

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

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

1 (PDF) ARMY RSCH LAB – HRED
HUMAN SYSTEMS INTEGRATION ENGR
TACOM FIELD ELEMENT
RDRL HRM CU P MUNYA
6501 E 11 MILE RD
MS 284 BLDG 200A
WARREN MI 48397-5000

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

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

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

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

1 (PDF)	ARMY RSCH LAB – HRED HQ USASOC RDRL HRM CN R SPENCER BLDG E2929 DESERT STORM DR FORT BRAGG NC 28310
1 (PDF)	ARMY G1 DAPE MR B KNAPP 300 ARMY PENTAGON RM 2C489 WASHINGTON DC 20310-0300

ABERDEEN PROVING GROUND

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

NTENTIONALLY LEFT BLANK