NORTH ATLANTIC TREATY ORGANIZATION

SCIENCE AND TECHNOLOGY ORGANIZATION

AC/323(HFM-237)TP/764

www.sto.nato.int

STO TECHNICAL REPORT **TR-HFM-237**

Assessment of Intelligent Tutoring Systems Technologies and Opportunities

(Evaluation et opportunités des technologies des systèmes de tutorat intelligents)

This Report documents the findings of the Human Factors and Medicine Task Group 237.

Published January 2018

Distribution and Availability on Back Cover

NORTH ATLANTIC TREATY
ORGANIZATION

SCIENCE AND TECHNOLOGY
ORGANIZATION

AC/323(HFM-237)TP/764

www.sto.nato.int

STO TECHNICAL REPORT **TR-HFM-237**

Assessment of Intelligent Tutoring Systems Technologies and Opportunities

(Evaluation et opportunités des technologies
des systèmes de tutorat intelligents)

This Report documents the findings of the Human
Factors and Medicine Task Group 237.

The NATO Science and Technology Organization

Science & Technology (S&T) in the NATO context is defined as the selective and rigorous generation and application of state-of-the-art, validated knowledge for defence and security purposes. S&T activities embrace scientific research, technology development, transition, application and field-testing, experimentation and a range of related scientific activities that include systems engineering, operational research and analysis, synthesis, integration and validation of knowledge derived through the scientific method.

In NATO, S&T is addressed using different business models, namely a collaborative business model where NATO provides a forum where NATO Nations and partner Nations elect to use their national resources to define, conduct and promote cooperative research and information exchange, and secondly an in-house delivery business model where S&T activities are conducted in a NATO dedicated executive body, having its own personnel, capabilities and infrastructure.

The mission of the NATO Science & Technology Organization (STO) is to help position the Nations' and NATO's S&T investments as a strategic enabler of the knowledge and technology advantage for the defence and security posture of NATO Nations and partner Nations, by conducting and promoting S&T activities that augment and leverage the capabilities and programmes of the Alliance, of the NATO Nations and the partner Nations, in support of NATO's objectives, and contributing to NATO's ability to enable and influence security and defence related capability development and threat mitigation in NATO Nations and partner Nations, in accordance with NATO policies.

The total spectrum of this collaborative effort is addressed by six Technical Panels who manage a wide range of scientific research activities, a Group specialising in modelling and simulation, plus a Committee dedicated to supporting the information management needs of the organization.

- AVT Applied Vehicle Technology Panel
- HFM Human Factors and Medicine Panel
- IST Information Systems Technology Panel
- NMSG NATO Modelling and Simulation Group
- SAS System Analysis and Studies Panel
- SCI Systems Concepts and Integration Panel
- SET Sensors and Electronics Technology Panel

These Panels and Group are the power-house of the collaborative model and are made up of national representatives as well as recognised world-class scientists, engineers and information specialists. In addition to providing critical technical oversight, they also provide a communication link to military users and other NATO bodies.

The scientific and technological work is carried out by Technical Teams, created under one or more of these eight bodies, for specific research activities which have a defined duration. These research activities can take a variety of forms, including Task Groups, Workshops, Symposia, Specialists' Meetings, Lecture Series and Technical Courses.

The content of this publication has been reproduced directly from material supplied by STO or the authors.

Published January 2018

ISBN 978-92-837-2091-1

Table of Contents

List of Figures and Tables

HFM-237 Membership List

CO-CHAIRMEN

Robert A. SOTTILARE
US Army Research Laboratory
12423 Research Parkway
Orlando, FL 32826
UNITED STATES
Email: Robert.A.Sottilare.civ@mail.mil

Ray S. PEREZ
US Office of Naval Research
875 N Randolph Street
Arlington, VA 22203-1995
UNITED STATES
Email: Ray.Perez@navy.mil

MEMBERS

Thomas ALEXANDER
Fraunhofer-FKIE
Zanderstr. 5
D-53177 Bonn
GERMANY
Email: Thomas.Alexander@fkie.fraunhofer.de

Scott A. DOUGLASS
US Air Force Research Laboratory
2210 8th Street Building 146
Wright-Patterson AFB, OH 45433-7511
UNITED STATES
Email: scott.douglass.1@us.af.mil

Paula J. DURLACH
US Army Research Laboratory
12423 Research Parkway
Orlando, FL 32826
UNITED STATES
Email: Paula.J.Durlach.civ@mail.mil

John D. FLETCHER
Institute for Defense Analyses
4850 Mark Center Drive
Alexandria, VA 22311-1882
UNITED STATES
Email: Fletcher@ida.org

Andrew P. GREEN
Defence Science and Technology
Laboratory (Dstl)
Portsdown West
Fareham, Hampshire PO17 6AD
UNITED KINGDOM
Email: apgreen@dstl.gov.uk

Pia JUSTEN
National Aerospace Laboratory (NLR)
Anthony Fokkerweg 2
1059 CM Amsterdam
NETHERLANDS
Email: Pia.Justen@nlr.nl

Paul NOJA
Gajon Institute of Technology
Via Paolo della Cella, 13
16135 Genova
ITALY
Email: paulnoja@gmail.com

Susan M. SIDEY
Defence Science and Technology
Laboratory (Dstl)
Portsdown West
Fareham, Hampshire PO17 6AD
UNITED KINGDOM
Email: smsidey@dstl.gov.uk

Vladimir ZOTOV
Defence R&D Canada – Toronto
1133 Sheppard Ave. West
Toronto, Ontario M3K 2C9
CANADA
Email: Vladimir.Zotov@drdc-rddc.gc.ca

CONTRIBUTING AUTHORS

Thomas ALEXANDER
Fraunhofer-FKIE
Zanderstr. 5
D-53177 Bonn
GERMANY
Email: Thomas.Alexander@fkie.fraunhofer.de

Paula J. DURLACH
US Army Research Laboratory
12423 Research Parkway
Orlando, FL 32826
UNITED STATES
Email: Paula.J.Durlach.civ@mail.mil

John. D. FLETCHER
Institute for Defense Analyses
4850 Mark Center Drive
Alexandria, VA 22311
UNITED STATES
Email: Fletcher@ida.org

Ray S. PEREZ
US Office of Naval Research
875 N Randolph Street
Arlington, VA 22203-1995
UNITED STATES
Email: Ray.Perez@navy.mil

Anna SKINNER
Black Moon, LLC
Washington, DC 20901
UNITED STATES
Email: Anna.Skinner@atinc.com

Robert A. SOTTILARE
US Army Research Laboratory
12423 Research Parkway
Orlando, FL 32826
UNITED STATES
Email: Robert.A.Sottilare.civ@mail.mil

Vladimir ZOTOV
Defence R&D Canada – Toronto
1133 Sheppard Ave. West
Toronto, Ontario M3K 2C9
CANADA
Email: Vladimir.Zotov@drdc-rddc.gc.ca

TASK GROUP MENTOR

Paul CHATELIER
US Naval Postgraduate School
8021 W. Point Drive
Springfield, VA, 22153
UNITED STATES
Email: Paul.R.Chatelier.civ@mail.mil

Assessment of Intelligent Tutoring Systems Technologies and Opportunities

(STO-TR-HFM-237)

Executive Summary

People ... the most critical elements of any military system are well-trained soldiers, sailors, airmen or marines. Present and future military missions conducted by NATO or their member countries will require highly trained individuals and collective teams to perform in extreme environments across a wide variety of task domains. It is critical that NATO exploit every tool and method available to assure a trained and ready military. Traditional classroom training is one method used in most NATO countries, but it is much less effective than one-to-one human tutoring. Ideally, NATO countries would provide tutors for every military member, but it is not practical to provide one-to-one human tutoring in every task domain required by NATO. The NATO Training Group's (NTG) working group on Individual Training and Educational Development (IT&ED) identified substantial instructional efficiencies to be achievable through the use of computer technology. Opportunities identified included both reduced costs and enhanced training effectiveness. However, most of the training effects were limited to memorization, understanding, and application of relatively straightforward facts, concepts, and procedures, and failed to exercise higher level cognitive skills (e.g., problem-solving and decision-making), provided few options to exercise skills in psychomotor domains (e.g., marksmanship), and almost no options to exercise collective or collaborative (social/team) skills.

Recently, Intelligent Tutoring Systems (ITSs) have begun to show equivalent effects to expert human tutors in providing tailored or adaptive learning experiences, but these experiences were primarily in well-defined cognitive domains. Still, the promise of adaptive instruction provided by computers might be a viable training option if the primary challenges could be identified and solutions discovered to enhance and accelerate learning. The recommendation of Research Task Group (RTG) HFM-165 in their final report along with the proposed work plan by HFM Exploratory Team (ET)-120 prompted NATO to charter HFM-237 Research Task Group (RTG) to investigate both existing and emerging ITS technologies and identify opportunities for their use in NATO training. This report summarizes the RTG's work plan and their findings along with chapters on ITS processes and applications. The report is a ready reference of ITS research and technology. The RTG examined the literature and activities in countries within and outside of NATO to discover the background, opportunities, and limits of ITS technologies (tools and methods). Tailoring experiences to individual learners and teams is the key to enhancing learning effect and accelerating the pace of learning.

Review of the literature, current and emerging research, and prototype development identified challenges in four major areas: authoring (development), standardization, data analytics, and adaptive interfaces. ITSs are currently expensive to author and require specialized skills including domain knowledge, instructional design, and computer programming. Emerging technologies are beginning to reduce the authoring burden through automation and enhanced usability. Authoring tools have been used to develop prototype tutors for both psychomotor and collective task domains which are increasing the relevance and ROI for ITS technologies with respect to military training needs. Standards may be another strategy to reduce the cost of ITSs by increasing reuse of ITSs and their components by promoting interoperability between learner models, standardizing instructional strategies, developing common communication protocols, and leveraging existing standards (e.g., experience Application Programming Interface – xAPI) to model learner competencies in various task domains. The ability to understand and model our learner population, course content, and instructional strategies through data analytics will allow training developers and managers to adapt ITS technologies to optimize learning outcomes. Finally, adaptive interfaces should be dynamic to adapt to the needs of individuals and their varying roles in ITS development, deployment, and evaluation.

Evaluation et opportunités des technologies des systèmes de tutorat intelligents

(STO-TR-HFM-237)

Synthèse

Les éléments essentiels de tout système militaire reposent sur des gens bien formés : soldats, marins ou aviateurs. Les missions militaires actuelles et futures menées par l'OTAN ou ses pays membres ont besoin de personnes très entraînées et d'équipes soudées, capables d'intervenir dans des environnements extrêmes et d'accomplir une grande diversité de tâches. Il est fondamental que l'OTAN exploite chaque outil et méthode à sa disposition pour garantir l'entraînement et l'état de préparation de ses militaires. La formation traditionnelle en classe est une méthode employée dans la plupart des pays de l'OTAN, mais elle est beaucoup moins efficace qu'un tutorat humain en tête-à-tête. Dans l'idéal, les pays de l'OTAN fourniraient des tuteurs à chaque membre de leur armée, mais cela n'est pas possible dans tous les domaines d'activité requis par l'OTAN. Le Groupe OTAN d'entraînement (NTG) sur la formation individuelle et le perfectionnement pédagogique (*IT&ED, Individual Training and Educational Development*) a identifié d'importantes améliorations pédagogiques pouvant être obtenues par l'informatique. Les opportunités d'amélioration incluaient à la fois la réduction des coûts et le renforcement de l'efficacité de la formation. Néanmoins, la plupart des effets de la formation se limitaient à la mémorisation, la compréhension et l'application de faits, concepts et procédures relativement simples ; la formation n'est pas parvenue à exercer les compétences cognitives plus complexes (telles que la résolution de problèmes et la prise de décision), proposait peu d'options pour exercer les compétences psychomotrices (par exemple, l'adresse au tir) et presque aucune option pour exercer des compétences collectives ou collaboratives (sociales / en équipe).

Récemment, des systèmes de tutorat intelligent (*ITSs, Intelligent Tutoring Systems*) ont commencé à produire des effets équivalents à ceux de tuteurs humains spécialisés, du point de vue de la pédagogie personnalisée ou adaptative, mais ces expériences concernaient principalement des domaines cognitifs bien définis. Un enseignement adaptatif délivré par des ordinateurs pourrait cependant être une option de formation viable, à condition d'en identifier les principaux obstacles et de trouver des solutions pour faciliter et accélérer l'apprentissage. Dans son rapport final, outre le plan de travail proposé par l'équipe exploratoire ET-120 de la HFM, le groupe de recherche RTG HFM-165 de la Commission HFM, incite l'OTAN à missionner le RTG HFM-237 de la HFM afin d'étudier les technologies ITS existantes et émergentes et à identifier leurs opportunités d'utilisation dans la formation OTAN. Ce rapport résume le plan de travail du RTG et ses conclusions et consacre un chapitre aux processus ITS et un autre à leurs applications. Il constitue un document de référence facilement consultable au sujet des recherches et technologies ITS. Le RTG a examiné la littérature et les activités de pays faisant ou non partie de l'OTAN, afin de découvrir le contexte, les opportunités et les limites des technologies ITS (outils et méthodes). L'adaptation de la formation à chaque individu et à chaque équipe est la clé si l'on veut renforcer l'effet de l'apprentissage et en accélérer le rythme.

La revue de la littérature, des recherches actuelles et émergentes et du développement des prototypes a révélé quatre champs principaux : la conception (développement), la normalisation, l'analyse des données et les interfaces adaptatives. Les ITS sont actuellement coûteuses à concevoir et nécessitent des compétences spécialisées qui incluent la connaissance du domaine, la conception du matériel pédagogique et la programmation informatique. Les technologies émergentes commencent à réduire cette charge de travail à travers l'automatisation et l'amélioration de la convivialité. Les outils de conception servent à développer des prototypes de tuteurs – tant dans le domaine psychomoteur que celui des tâches collectives – qui améliorent la pertinence et la rentabilité des technologies ITS concernant la formation militaire. Les normes peuvent être une autre stratégie de réduction du coût des ITS, puisqu'elles augmentent la réutilisation des ITS et de leurs composantes : elles favorisent l'interopérabilité entre les modèles d'apprentissage, normalisent les stratégies pédagogiques, développent des protocoles de communication communs et exploitent les normes existantes (par exemple, l'interface applicative d'expérience, xAPI) pour modéliser les compétences des élèves dans divers domaines. La capacité à comprendre et modéliser notre population d'élèves, le contenu des cours et les stratégies pédagogiques, et ce, grâce à l'analyse des données, permettra aux développeurs et aux responsables des formations d'adapter les technologies ITS pour optimiser l'apprentissage. Enfin, les interfaces adaptatives devraient être dynamiques et s'adapter aux besoins des personnes et à leurs rôles variables en ce qui concerne le développement, la mise en place et l'évaluation des ITS.

Chapter 1 − RESEARCH TASK GROUP (HFM-237) WORK PLAN

Robert A. Sottilare, Ph.D. and J.D. Fletcher, Ph.D.

This chapter summarizes the background and justification for Research Task Group (RTG) HFM-237, *Assessment of Intelligent Tutoring System (ITS) Technologies and Opportunities*, for the period March 2013 – August 2016. It describes the objectives, provides definitions, and lists topics covered by RTG HFM-237. Later chapters summarize the background of intelligent tutoring systems, provide reviews of ITS authoring tools, discuss various findings concerning emerging ITS technologies, discuss applications of ITSs for collective (team) training, examine the topic of education in Science, Technology, Engineering, and Mathematics (STEM), review the application of ITS to medical education and training, and, in a final wrap-up chapter, review overall capabilities, gaps, challenges, and recommendations for moving ITS technologies from state-of-art to state-of-practice.

1.1 JUSTIFICATION FOR RTG HFM-237

The NATO Training Group's (NTG) working group on Individual Training and Educational Development (IT&ED) found substantial instructional efficiencies in terms of both reduced costs and enhanced effectiveness to be achievable through the use of computer technology. However, most of these effects concerned memorization, understanding, and application of relatively straightforward facts, concepts, and procedures as early reported by Vinsonhaler and Bass [1] and later confirmed by Kulik [2], who surveyed 97 studies of basic computer-based instruction effectiveness for his 1994 meta-analysis and found an average effect size of 0.32 standard deviations (σ), which is roughly equivalent to novice human tutoring. Although the capabilities produced by such basic instruction are vital to successful accomplishment of military operations, they are not sufficient.

Military operations, especially those characteristic of current irregular warfare environments, require, among other things, improvisation, rapid judgment, and the ability to deal with the unexpected. They go beyond basic instructional objectives and call for education and training focused on higher order cognitive capabilities such as analysis, evaluation, creativity, and rapid synthesis of novel approaches – approaches that must intersperse judgment with the automatic responses provided by training involving memorization and practice of straightforward procedures. These capabilities can make the difference between success and failure in operations, and require more sophisticated forms of instruction such as one-to-one tutoring.

While it is not practical to provide one-to-one human tutoring to every soldier, sailor, and airman, it is practical to provide computer-based tutoring that is dynamically tailored to every learner's capabilities and needs. Technology to produce such tutoring requires 'intelligent' systems that rapidly tailor instruction to individual learner abilities, prior knowledge, experience, and, to some extent, misconceptions (e.g., common errors or malformed mental models about the training domain). As with basic instruction, technology is required to make this education and training practical, affordable, and effective at the very large scale required for military personnel. This technology has long been the goal of approaches earlier labeled Intelligent Computer Assisted Instruction (ICAI) and, more recently, Intelligent Tutoring Systems (ITSs).

Research at universities (in Europe, North America, and elsewhere) and some commercial enterprises has produced effect sizes in excess of 1.00 σ (roughly an increase of learner performance from the 50^{th} to the 84^{th} percentile) and, occasionally, 2.00 σ (roughly an increase from the 50^{th} percentile to the 98^{th}). Recent research on digital tutoring by the US Defense Advanced Research Projects agency (DARPA) has found effect sizes of 3.00 σ and 4.00 σ for one of their ITS efforts [3], but this is not the norm. Recent reviews of ITSs have found effect sizes averaging 0.70 to 0.80 σ or about the equivalent of one letter grade primarily, but not solely, in well-defined cognitive domains such as computer programming, mathematics, and physics [4], [5]. These results are similar to those produced by highly competent one-to-one human tutors.

The promise of significant growth in the effectiveness of ITSs over the next 5 to 10 years is very real. ITSs may also offer NATO substantial and unique opportunities for developing instruction that develops the critical cognitive readiness capabilities needed for military operations. Application of ITS technology in task, job, and performance aids may additionally contribute to mission success. The nature, extent, availability, and feasibility of these opportunities was identified, reviewed, and assessed for NATO applications by members of RTG HFM-237.

1.2 RTG HFM-237 OBJECTIVES

The RTG reviewed and analyzed the nature, extent, availability, and feasibility of opportunities presented by ITSs for conducting NATO education and training. While their effectiveness as one-to-one instructional tools is well documented, several barriers are challenging their widespread adoption in military training and education:

- Highly specialized skills are required to build effective ITSs – deep understanding of instructional design, one-to-one tutoring, the subject matter, and computer programming is required so the entry level to authoring ITSs is currently high.
- ITSs are expensive and time consuming to build – this limits return-on-investment for ITSs in complex domains with low density training (low throughput) low.
- The more adaptive an ITS is, the more content it needs to support tailoring and personalization of instruction – which also leads to longer development times and higher costs.
- Interoperability and data exchange between tutoring systems is severely limited by the lack of ITS standards for reuse, portability, and sharing of content, for models of learners, pedagogy, and domains, and protocols such as those for internal and external messaging.

Overcoming these challenges is essential for ITSs and their associated adaptive instructional techniques to become practical for use in large scale organizations (e.g., military schools and training centers). Each of these challenges is addressed in our recommendations in Chapter 8.

It is essential to note that ITSs are technology-based solutions for adaptive instruction in specified domains. Their practicality for use in education and training in all sectors depends on:

- Their ability to accurately model the learner and the training environment (e.g., a set of problems or immersive virtual environment).
- Their ability to effectively adapt instructional interaction to the learner's capabilities and needs.
- The likelihood that existing ITS technologies (tools and methods) may be reused and extended to support military training domains.
- The skills and cost required to author, distribute, apply, and evaluate their effectiveness.

Based on this goal of practicality, RTG-237 members canvassed adaptive training and education research and development efforts around the world to more fully understand the antecedents of effective and affordable ITSs.

1.3 INTELLIGENT TUTORING SYSTEMS

This section defines ITSs, as used by RTG HFM-237 in this report and elsewhere, in order to compare and contrast them with more conventional computer-based instructional technologies such as computer-Aided Instruction (CAI) and Computer-Based Training (CBT). The definition of ITSs varies across researchers, designers, and developers. According to Fletcher and Sottilare [6], *intelligent tutoring* may be viewed as "an effort to capture in computer technology the capabilities and practices of a human instructor who is expert in both the subject matter and one-to-one tutoring". Intelligent tutoring is a form of *adaptive instruction*. The technology to deliver and manage adaptive instruction is critical for ITS.

ITS development is motivated by the empirically evident benefits of human tutoring [1], [6]-[7] and a long standing desire to make these benefits more widely accessible and affordable than those delivered by human tutors [8]-[11].

Another motivation for the development of ITSs grew from the recognition that although computers could be used to teach effectively, it took time, effort, and considerable expense to anticipate all possible states of the learner and to program all possible instructional responses to these states. Response to both of these issues requires a generative capability, which is a defining characteristic of ITS [12]. Dynamic information structures and mixed-initiative computer-based tutorial dialogue are required to generate instructional interactions in real time thereby relieving much of the burden and cost of authoring adaptive, individualized instruction [8], [12]-[13]. Not only must ITSs be more adaptive and responsive to learners than simple CBT, they must eventually become self-authoring systems. With the capability to access almost all human knowledge through the global information infrastructure, ITS capabilities may make learning affordable and universally accessible, generated on demand − anytime and anywhere [8], [14].

Today ITS development suggests a future in which education, training, and performance aiding do not take place solely through prefabricated lessons and other material but are provided in the form of one-to-one, guided dialogues, that are generated on demand, tailored to the needs, abilities, interests, and values of individual learners, and are based on mixed-initiative conversations in which either the computer/tutor or the learner may take the initiative [14].

ITSs can be contrasted with drill and practice programs. The latter methods were found to be very effective in achieving lower level instructional objectives such as learning arithmetic facts [15], grapheme-phoneme correspondences in beginning reading [16], and foreign language vocabulary and phonetics [17].

The rudimentary objectives of drill and practice are found in initial learning of nearly all subject domains. They consist of discrete items, simple concepts, nomenclature, or straightforward procedures to be memorized and/or applied and are limited to objectives in the lower skill of Bloom's [18] hierarchy or the lower left-hand corner of Anderson and Krathwohl's 2-dimension learning space [19]. Drill and practice programs have a strong role to play at this level. They are effective and inexpensive to design, develop, and deliver. They require models of the learner, but all relevant states of the learner must be anticipated at design time and pre-programmed into the system. Learner modeling in these systems is predominately pre-assigned, implicit, and categorical. As effective as drill and practice programs are for helping learners master domain rudiments, they are limited when compared to the higher level objectives that are properly targeted by ITSs.

ITS are not unique in their use of learner models, but their approach to learner modeling substantially augments and extends those needed for drill and practice [20], [21], [22]. ITS learner models are dynamic and may be generated on demand as needed by the instructional program. They are based on explicit, comprehensive models of the procedures, knowledge, and skills required to progress toward instructional objectives. Because of their dynamic qualities, they are particularly suited to tutorial dialogue systems that generate instructional and problem solving guidance on demand, in real time.

In recent years, military trainers have begun to push for intelligent systems that support training and education for both individuals and collectives such as crews, teams, and units. These collectives, which are the building blocks of military organizations, are essential in meeting challenges associated with organizational missions. Collective training is often the norm as programs of instruction strive to align with operational needs. The benefits of guided instruction for collectives may be similar to individual adaptive instruction. While insufficient attention and research resources have been focused on training and education for collectives [3], [22], [23], collective tutoring and collaborative problem solving [24], [25] have historically received and will continue to require focused attention from military organizations. Research to address these issues should be strongly encouraged and pursued to ensure the availability of essential capabilities for training collectives. Emerging ITS techniques have much to contribute in this area by modeling the states of collectives, improving their capabilities, and substantially enhancing their mission effectiveness.

1.4 TOPICS COVERED BY RTG HFM-237

Following the discussion on the definition and scope covered by ITSs, this section reviews the topic areas covered under the investigation conducted by RTG HFM-237. In addition to the technical topic areas covered by the RTG, it also discussed and documented Return-On-Investment (ROI) considerations in Chapter 2 of this Report. The technical topic areas pursued by the Group were primarily focused on task domains and processes:

a) Authoring tools (Chapter 3);

b) Emerging ITS technologies (Chapter 4);

c) Collective or team-based tasks (Chapter 5);

d) Science, Technology, Engineering, and Mathematics (STEM) tasks (Chapter 6); and

e) Medical tasks (Chapter 7).

A summary of each topic area follows. The individual chapters provide additional details later in this Report.

1.4.1 Authoring Tools

While authoring tools was not a theme for any of the RTG's formal meetings, it was a topic of interest based on the high cost and skill required to develop ITSs. A variety of authoring tools are associated with the process of developing the primary elements of ITSs which usually include learner, instructional and domain models along with a tutor-user interface to facilitate the capture of physiological or behavioral data captured by sensors and learner input. These tools may include default policies or learning strategies which capture domain-independent best practices or dashboards to allow developers to link learner states and traits to specific actions by the tutor. In an effort to reduce the authoring workload, training scientists are also examining the use of artificial intelligence to automatically generate content, plan tutorial routes, and adapt or evolve multiple scenarios from a single parent scenario.

1.4.2 Emerging ITS Technologies

In discussing emerging technologies, we begin with existing systems the research investments being made that will yield new capabilities in the next few years. The list of existing systems that follows is intended to provide a sampling of capabilities evolving from ITS research. As noted previously, ITSs primarily focus on well-defined domains (computer programming, mathematics, and physics).

1.4.2.1 Shell (Multi-Domain) Tutors

Generally, each ITS is custom-built to support training in a single domain, but there are multi-domain tutoring architectures also known as shell tutors. Three of these are discussed below:

- The *Generalized Intelligent Framework for Tutoring (GIFT)*, developed by the Learning in Intelligent Tutoring Environments (LITE) Lab at the US Army Research Laboratory, is emerging as a multi-domain, open source tutoring architecture [26]. GIFT is a research prototype intended to reduce the computer skills and cost required to author ITSs, deploy them, manage them, and continuously evaluate the adaptive instruction they provide. A major advantage of GIFT is that three of its four functional elements are reusable across task domains. GIFT may also be linked to external training environments (e.g., serious games or virtual and augmented reality simulations) through a standardized gateway. GIFT authoring tools require no knowledge of computer programming or instructional design to develop effective ITSs. GIFT is freely available and may be hosted either locally or cloud-based. GIFT-based tutors have been prototyped to support training in adaptive marksmanship, land navigation, medical casualty care, and other military and non-military domains. GIFT, like other ITS technologies, has focused on training individuals, but research is underway to create tools and methods to support tutoring of collectives. At the time of this writing, GIFT has a community of over 800 government, academic, and industry users in 53 countries. Additional information about GIFT is available at www.GIFTtutoring.org.
- *AutoTutor*, developed at the University of Memphis, has been a stalwart in dialogue-based tutoring over the last 20 years. AutoTutor is an intelligent tutoring system that holds conversations with the human learner in natural language. AutoTutor has produced learning gains across multiple domains (e.g., computer literacy, physics, critical thinking). AutoTutor research is focused on three main areas: human-inspired tutoring strategies, pedagogical agents, and natural language tutoring. AutoTutor has been applied to several task domains in support of one-to-one tutoring, and it has a comprehensive set of authoring tools and services. An emerging capability in AutoTutor is the *trialogue*, intelligent pedagogical agents that help students learn by holding a conversation in natural language between the student, a virtual instructor, and a virtual student [27]. Additional information about AutoTutor is available at www.autotutor.org/.
- *ASPIRE*, developed by the University of Canterbury in New Zealand, is a system for developing and delivering adaptive instruction on the web. The system consists of ASPIRE-Author, a tutor development server, and ASPIRE-Tutor, a tutoring server that delivers the resulting ITSs to students for guided instruction. The authoring system provides a unique process for composing an ontology of the domain by outlining basic domain concepts, their properties, and the relationships between concepts forming the basis of an expert model. Lessons learned from the ASPIRE authoring process may reduce the time and cost associated with authoring ITSs and/or increase the accuracy of the represented domain. Additional information about ASPIRE is available at http://aspire.cosc.canterbury.ac.nz/.

1.4.2.2 Single Domain Tools

The HFM-237 Task Group also examined single domain ITS tools and methods to identify potential transition opportunities to military domains. Several technologies, described below, are noteworthy candidates for application in NATO training and education:

- *DynaLearn*, developed at the University of Amsterdam, is an intelligent environment for teaching conceptual knowledge, which is viewed as a connected web of knowledge that cannot be learned through rote memorization, but must be learned through thoughtful reflection. DynaLearn allows learners to acquire conceptual knowledge by constructing and simulating qualitative models of how physical systems behave. Modeling has been considered fundamental to human cognition and scientific inquiry [28] as way of obtaining a deep understanding of a scientific domain knowledge structure and content. The DynaLearn approach is based on the hypothesis that building a conceptual model helps student acquire a better understanding of scientific topics. Additional information about DynaLearn is available at: https://ivi.fnwi.uva.nl/tcs/QRgroup/DynaLearn/.

- The *Adaptive Courseware Tutor (AC-ware Tutor)*, developed at the University of Split in Croatia, includes automatic generation of courseware elements, dynamic selection and sequencing of courseware elements, automatic generation of tests and questions, including an initial test over a representative subset of the domain knowledge. This tutor adapts to changes in the learner's knowledge. While AC-ware Tutor is focused on the mathematics domain, the process and logic in AC-ware Tutor might be applied to future military ITSs. Additional information about AC-ware Tutor is available at: http://ijitcs.com/volume%2016_No_1/Volume_16 No_1.php.

- The *Basic Electricity and Electronics Learning Environment (BEETLE II)*, developed under the direction of the US Naval Warfare Center's Training Systems Division and jointly with the University of Edinburgh, has advanced computer-based tutorial dialogue capabilities to support conceptual learning in electricity and electronics. BEETLE II is capable of interpreting student natural language explanations to support active experimentation, self-explanation, and generation principles. This tutor provides relevant one-to-one tutoring in a military domain, and opportunities to reuse its tutoring processes are considered significant. Additional information about BEETLE II is available at: http://groups.inf.ed.ac.uk/beetle/beetle_bib.html.

- *The DARPA Digital Tutor*, was developed to accelerate the acquisition of expertise in Information systems Technology (IT) by novice US Navy sailors. The DARPA challenge was to capture in computer technology the practice and capabilities of individuals who were established experts both in relevant sub-areas of IT and in one-to-one tutoring. The Tutor is a problem-based dialogue system. It uses information structures and models derived from experts as they tutored learners similar to new sailors. It uses these structures and models to provide a problem-based, tutorial environment similar to that provided by the human tutors. This approach enables the Tutor to match individual learners with relevant tutoring based on the knowledge earlier collected from experts. It provides brief didactic information, which is immediately followed by progressively more difficult problem-solving exercises intended to accelerate acquisition of the deeper, conceptual levels of knowledge that considerable research and analyses has found to be essential for retention and transfer of learning. It provides a 16-week course of instruction that was shown to graduate ITs with the knowledge and skills of those who have expertise gained from an average of 9 years of Fleet experience [3]. Additional information about the DARPA Tutor is available at: http://www.acuitus.com/web/education-dominance.html.

- *QuestionIT*, created by Defence Research & Development Canada, guides members of the Canadian Armed Forces in acquiring skills to effectively question witnesses following Improvised Explosive Device (IED) attacks. QuestionIT guides learning through a strategy of *worked examples*, a step-by-step solution to a problem or task. The ability to efficiently implement worked examples, especially faded worked examples, during ITS authoring is highly desirable. For faded worked examples [29], initially, all the worked-out steps are presented together with self-explanation prompts. As the learner demonstrates understanding, the worked-out steps are faded − gradually withdrawn. They begin with examples where the complete solution is presented to the learner and evolve to problems where learner must find the whole solution. Additional information about QuestionIT is available at: http://cradpdf.drdc-rddc.gc.ca/PDFS/unc194/p801512_A1b.pdf.

The next subsections examine intelligent tutoring/adaptive instruction within three relevant task domains.

1.4.3 Intelligent Tutoring of Collective or Team-Based Tasks

In recent years, military trainers [30] have begun to recognize that ITSs have the potential to support training and education of both individuals and teams. Teams, the basic building blocks of most organizations, are important in meeting the challenges associated with accomplishing organizational missions. Team or collective training is the norm in military organizations. There is no denying the importance of improving individual competence, but managers and commanders at all levels generally focus on the collective performance of the teams for which they are responsible and accountable. Given the potential of ITSs and the importance of teams, it is astonishing that so little attention or resources are provided to identify the capabilities required to support intelligent tutoring of teams [31]. One area where there has been a recent surge in research investments is in tutoring teams during collaborative problem solving [32], [33], [34]. Lessons from team tutoring research efforts at the US Army Research Laboratory and in collaborative problem solving are expected to yield insights into the development of ITSs for teams. Combining ITS experience in modeling learners and their learning progress with the evolving understanding of the shared mental models developed by members of collectives should provide a productive area of research and development for ITS applied to the training of collectives [6], Chapter 5 in this report provides more comprehensive discussion of team tutors.

1.4.4 Intelligent Tutoring of STEM Tasks

The tutoring of STEM-based tasks is important to the development of military personnel in science and engineering disciplines, but it may be equally important to military occupations that simply require learning to use new technology. The STEM domain is the most represented domain within deployed ITSs with mathematics and physics leading the way. Academic tutors for reading and writing (e.g., adult literacy tutors) are also prevalent. A prominent effort in the development of intelligent tutoring systems for STEM tasks has been sponsored by the US Office of Naval Research (ONR). Additional information about the ONR STEM Grand Challenge is available at: http://www.onr.navy.mil/Media-Center/Press-Releases/2011/STEM-grand-challenge-ONR.aspx. Chapter 6 in this Report provides more comprehensive discussion of STEM tutors.

1.4.5 Intelligent Tutoring of Medical Tasks

There is likely no set of military tasks where performance matters so much and where the opportunity for deliberate practice is so critical for proficiency as medical tasks [35]. The procedural nature of many medical tasks, fits seamlessly into the authoring paradigms of existing ITS architectures. A shortlist of relevant efforts related to the adaptive instruction of medical tasks includes a pathology ITS [36], “Docs ‘n drugs--the virtual polyclinic” (an intelligent tutoring system for web-based and case-oriented training in medicine) [37], several anatomy tutors, a physician’s assistant that is similar to a pilot’s assistant trainer, and a virtual medic tutor, which is a GIFT-based tutor linked to VMedic, a serious game to support tutoring of medical casualty care and triage tasks [38]. An example of an ITS that spans both medical and team tutoring domains is COMET, a collaborative ITS for medical problem-based learning [39]. The US Office of Naval Research is developing an ITS on the fundamentals of laparoscopic surgery. Progress in medical tutoring is substantial and transfer of technology from civilian to military application will likely be transparent due to the similarity of tasks and environments (e.g., triage at civilian hospital emergency rooms and mobile military hospitals). Chapter 7 in this Report provides more comprehensive discussion of medical tutors.

The next chapter in this Report provides additional background on ITS history, return on investment, human-machine interface issues, and instructional strategies.

1.5 REFERENCES

[1] Vinsonhaler, J.F. and Bass, R.K. (1972). "A Summary of Ten Major Studies on CAI Drill and Practice." *Educational Technology* 12: 29-32.

[2] Kulik, J.A. (1994). Meta-Analytic studies of findings on computer-based instruction. In E.L. Baker and H.F. O'Neil (Eds.), Technology Assessment in Education and Training (pp. 9-33). Hillsdale, NJ: Lawrence Erlbaum Associates.

[3] Fletcher, J.D., and Morrison, J.E. (2014). *Accelerating Development of Expertise: A Digital Tutor for Navy Technical Training* (D-5358). Alexandria, VA: Institute for Defense Analyses.

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

[5] Kulik, J.A., and Fletcher, J.D. (2016). Effectiveness of Intelligent Tutoring Systems: A Meta-analytic Review. *Review of Educational Research, 86,* 42-78.

[6] Fletcher, J.D. and Sottilare, R. (2013). Shared Mental Models of Cognition for Intelligent Tutoring of Teams. Design Recommendations for Intelligent Tutoring Systems: Volume 1- Learner Modeling. Army Research Laboratory, Orlando, Florida. ISBN 978-0-9893923-0-3.

[7] Bloom, B.S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. Educational Researcher, 13, 4-16.

[8] Fletcher, J.D. (1992). Individualized systems of instruction. In M.C. Alkin (Ed.) *Encyclopedia of Educational Research* (Sixth Edition) (pp. 613-620). New York, NY: Macmillan.

[9] Fletcher, J.D. (2009). Education and Training Technology in the Military. *Science*, *323*, 72-75.

[10] Corbett, A. (2001). Cognitive computer tutors: Solving the two-sigma problem. In M. Bauer, P.J. Gmytrasiewicz, and J. Vassileva (Eds.), *User modeling* (pp. 137-147). Berlin: Springer-Verlag.

[11] Sottilare, R. (2013). Adaptive Intelligent Tutoring System (ITS) Research in Support of the Army Learning Model - Research Outline. *Army Research Laboratory* (ARL-SR-0284), December 2013.

[12] Carbonell, J.R. (1970). AI in CAI: An Artificial Intelligence Approach to Computer-Assisted Instruction. *IEEE Transactions on Man-Machine Systems*, *11*, 190-202.

[13] Sottilare, R.A. and Holden, H.K. (2013). Motivations for a Generalized Intelligent Framework for Tutoring (GIFT) for Authoring, Instruction and Analysis. In *Proceedings of the 1st Generalized Intelligent Framework for Tutoring (GIFT) Users Symposium at the Artificial Intelligence in Education Conference* (AIED) 2013, Memphis, Tennessee, June 2013.

[14] Fletcher, J.D., Tobias, S. and Wisher, R.L. (2007). Learning Anytime, Anywhere: Advanced Distributed Learning and the Changing Face of Education. Educational Researcher, 36(2), 96-102.

[15] Suppes, P. and Morningstar, M. (1972). Computer-assisted instruction at Stanford 1966-68: Data, models, and evaluation of the arithmetic programs. New York: Academic Press.

[16] Fletcher, J.D. and Atkinson, R.C. (1972). An evaluation of the Stanford CAI program in initial reading (Grades K through 3). *Journal of Educational Psychology, 63,* 597-602.

[17] Van Campen, J. (1981). A computer-assisted course in Russian. In, P. Suppes (Ed.), *University-Level Computer-Assisted Instruction at Stanford: 1968-1980.* Stanford, CA: Institute for Mathematical Studies in the Social Sciences, Stanford University.

[18] Bloom, B.S. (1956). Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain. New York: David McKay Co. Inc.

[19] Anderson, L.W., Krathwohl, D.R., Airasian, P., Cruikshank, K., Mayer, R., Pintrich, P. and Wittrock, M. (2001). A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy. *New York. Longman Publishing. Artz, A.F. and Armour-Thomas, E.(1992). Development of a cognitive-metacognitive framework for protocol analysis of mathematical problem solving in small groups. Cognition and Instruction, 9*(2), 137-175.

[20] Corbett, A.T. and Anderson, J.R. (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction 4*(4), 253-278.

[21] Dütsch, I. and Gediga, G. (1995). Skills and knowledge structures. *British Journal of Mathematical and Statistical Psychology 48*, 9-27.

[22] Sottilare, R.A., Holden, H.K., Brawner, K.W. and Goldberg, B.S. (2011). Challenges and emerging concepts in the development of adaptive, computer-based tutoring systems for team training. *Interservice/Industry Training Systems & Education Conference*, Orlando, Florida, December 2011.

[23] Sottilare, R., Ragusa, C., Hoffman, M. and Goldberg, B. (2013). Characterizing an adaptive tutoring learning effect chain for individual and team tutoring. In *Proceedings of the Interservice/Industry Training Simulation and Education Conference*; 2013 December; Orlando, FL.

[24] Adamson, D., Dyke, G., Jang, H. and 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.

[25] Rosen, Y. (2015). Computer-based assessment of collaborative problem solving: Exploring the feasibility of human-to-agent approach. *International Journal of Artificial Intelligence in Education*, 25(3), 380-406.

[26] Sottilare, R.A., Brawner, K.W., Goldberg, B.S. and 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.

[27] Feng, S., Stewart, J., Clewley, D. and Graesser, A.C. (2015, June). Emotional, Epistemic, and Neutral Feedback in AutoTutor Trialogues to Improve Reading Comprehension. In International Conference on Artificial Intelligence in Education (pp. 570-573). Springer International Publishing.

[28] Schwarz, C.V. and White, B.Y. (2005). Metamodeling knowledge: Developing students' understanding of scientific modeling. *Cognition and instruction*, 23(2), 165-205.

[29] Salden, R.J., Aleven, V.A., Renkl, A. and Schwonke, R. (2009). Worked examples and tutored problem solving: redundant or synergistic forms of support? *Topics in Cognitive Science*, 1(1), 203-213.

[30] U.S. Army Training & Doctrine Command (TRADOC). (2011). The US Army Learning Concept for 2015 (525-8-2). Washington, DC: Department of the Army.

[31] 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.

[32] Adamson, D., Dyke, G., Jang, H. and 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.

[33] Rosen, Y. (2015). Computer-based assessment of collaborative problem solving: Exploring the feasibility of human-to-agent approach. *International Journal of Artificial Intelligence in Education*, 25(3), 380-406.

[34] McManus, M.M. and Aiken, R.M. (2016). Supporting Effective Collaboration: Using a Rearview Mirror to Look Forward. *International Journal of Artificial Intelligence in Education*, 26(1), 365-377.

[35] Ericsson, K.A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic medicine*, 79(10), S70-S81.

[36] Crowley R.S., Legowski, E., Medvedeva, O., Tseytlin, E., Roh, E. and Jukic, D. Evaluation of an Intelligent Tutoring System in Pathology: Effects of External Representation on Performance Gains, Metacognition, and Acceptance. *Journal of the American Medical Informatics Association* : JAMIA. 2007;14(2):182-190. doi:10.1197/jamia.M2241.

[37] Martens, A., Bernauer, J., Illmann, T. and Seitz, A. (2001). "Docs 'n drugs--the virtual polyclinic": an intelligent tutoring system for web-based and case-oriented training in medicine. In Proceedings of the AMIA Symposium (p. 433). *American Medical Informatics Association*.

[38] Pacquette, L., Rowe, J., Baker, R., Mott, B., Lester, J., DeFalco, J., Brawner, K., Sottilare, R. and Georgoulas, V. (2015). Sensor-Free or Sensor-Full: A Comparison of Data Modalities in Multi-Channel Affect Detection, *Electronic Data Mining (EDM) Conference*, Madrid, Spain, June 2015.

[39] Suebnukarn, S. and Haddawy, P. (2007). COMET: A collaborative tutoring system for medical problem-based learning. *IEEE Intelligent Systems*, *22*(4), 70-77.

Chapter 2 − A PERSPECTIVE OF INTELLIGENT TUTORING SYSTEMS DESIGN AND USE

J.D. Fletcher, Ph.D. and Robert A. Sottilare, Ph.D.

2.1 INTRODUCTION

It is evident and well confirmed by research that people differ in their abilities, interests, and values. Further, they differ in the rate with which they learn. Early research determined that the rate with which different students in a typical elementary or secondary (K-12) school classroom learn and it differs by a factor of at least 4:1 between high and low performers [1]. That finding suggests that an appreciable number of students in any classroom, which in the US averages about 25 students, may be bored, with their time to advance in the subject matter wasted while waiting for others to learn what is being taught. This problem was noted in 1906 by the psychologist E.L. Thorndike [2], who wrote that:

> "*The principal consequence of individual differences is that every general law of teaching has to be applied with consideration of the particular person . . . responses to any stimulus . . . will vary with individual capacities, interests, and previous experience*" (page 83).

The problem can be eased by some classroom practices, but only partially. It remains as an unavoidable bar to inefficiency and effectiveness in training and education using classroom approaches.

Both ability and prior knowledge affect learning rates. Although ability accounts for acquisition of knowledge, prior knowledge appears to be the most direct determinant of learning rates [3], [4]. The life experiences of older students, such as those in the military, are therefore likely to exceed the variety and quantity of those of younger, K-12 students. The disparities in their learning rates and prior knowledge increase the need for individualized instruction well beyond that of K-12 learners.

The problem is not that people in classrooms fail to learn, most of them obviously and demonstrably do. However, time and cost to produce an effective workforce are at a premium in military training. The inefficiencies and uneven results of classroom learning create severe monetary and operational problems. These problems may be viewed as inevitable, but research suggests some solutions.

2.2 TUTORING AND A ROLE FOR INTELLIGENT TUTORING SYSTEMS

One solution is to provide a human tutor for each learner. This is done for particular subjects, such as surgery and airplane piloting, where avoidance of errors is critical and the apprentice model is the norm. It is also used to produce especially high levels of expertise, which tend to be idiosyncratic and require intense individual attention. In nearly all cases, one-to-one tutoring has been far superior to one-on-many classroom instruction [5]-[7].

However effective it may be, one-on-one human tutoring is rarely affordable. Computers, with their rapid and accurate access to information, have, since their inception in the 1950s, motivated their application in training and education as a way to deal with individual differences. Computer-assisted instruction (CAI)[1] has, then, been

[1] A variety of labels are used to describe the use of computers to teach (e.g., Computer-Based Instruction, CBI; Computer-Aided Instruction, CAI; Computer-Aided Education, CAE; Computer-Assisted Learning, CAL; Computer-Assisted Instruction, CAI again). CAI is used here as a generic term for all applications of computers to the teaching-learning process.

sought as a way to provide every learner with an appreciable amount of the individualization needed for efficient and effective learning.

Figure 2-1 outlines a two-dimensional (content and objectives) learning space adapted from Anderson and Krathwohl [8]. It suggests a rough dichotomy of objectives with implications for how they might be learned most effectively and efficiently. Learning in most subjects begins with basic facts, nomenclature, and simple procedures that must be learned and, to an extent, memorized. Objectives for these essentials are found in the lower left-hand corner of the Anderson-Krathwohl learning space and are best presented and achieved using drill and practice as demonstrated by early research in CAI [9]-[13].

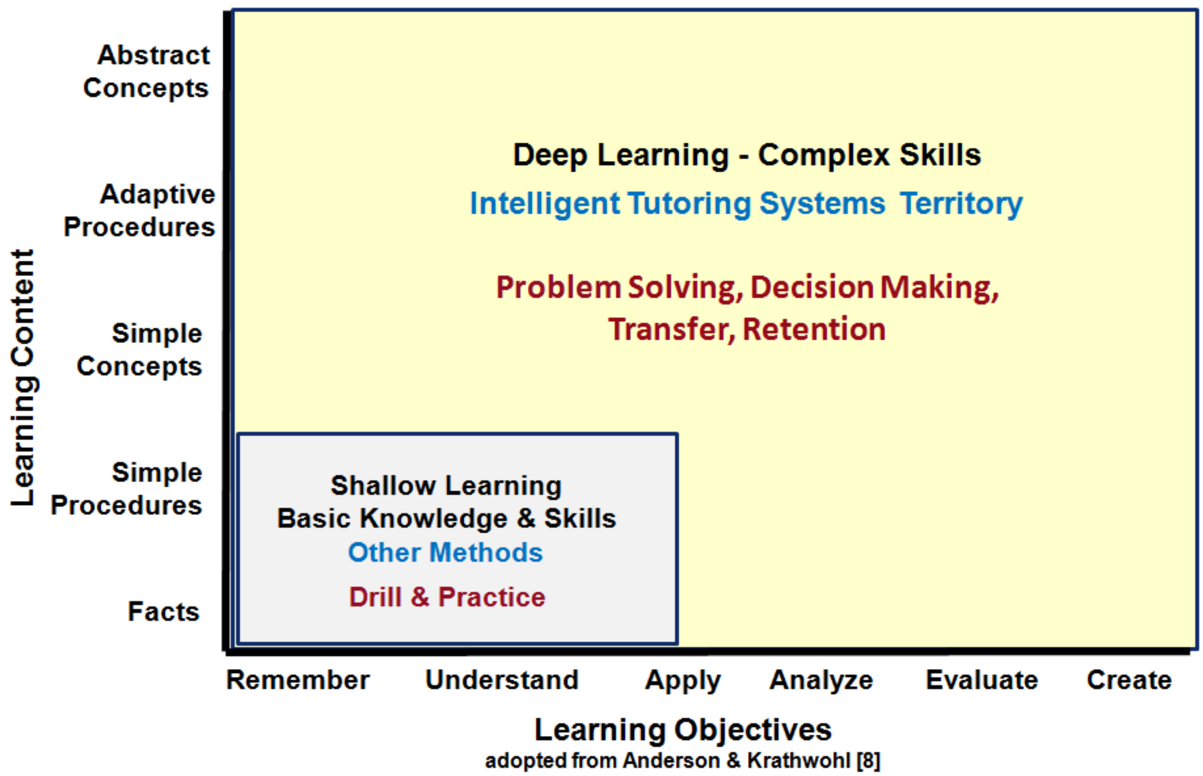

Figure 2-1: The Effective and Efficient Use of Intelligent Tutoring System Technologies.

These objectives are essential, but they do not constitute the deeper and more conceptual understanding that is required to advance from entry-level capabilities to the journeyman or expert levels needed for transfer and retention of knowledge [3], [6], [14]-[15]. Findings by these researchers and others, suggest that more relevant, persistent, and applicable material must be presented using more sophisticated techniques than drill and practice. Most of these techniques involve some form of one-on-one human tutoring, which is generally unaffordable unless it can be provided by computers that are imbued with some of the intelligence needed to clone human dialogue-based tutoring. The development of training and education based on machine or 'artificial' intelligence is, therefore, not driven by infatuation with technology or computers, but by a genuine operational need, if not an imperative, to provide a capability that is as effective and efficient as possible within the bounds and limits of resources commonly allocated to training and education.

2.3 COMPUTER-ASSISTED INSTRUCTION AND INTELLIGENT TUTORING SYSTEMS

The earliest techniques for CAI were adapted from techniques developed for programmed textbooks. These techniques remain in wide use today. They are illustrated in Figure 2-2, which shows Keller's [16] Personalized System of Instruction combined with Intrinsic Programming described and recommended for CAI by Crowder [17]. Keller's system separates instructional content into modules that are usually presented in a linear sequence. It requires students to demonstrate mastery of each unit before proceeding to the next. Once inside the module, Crowder's system takes over by emphasizing active responding and immediate feedback to learners.

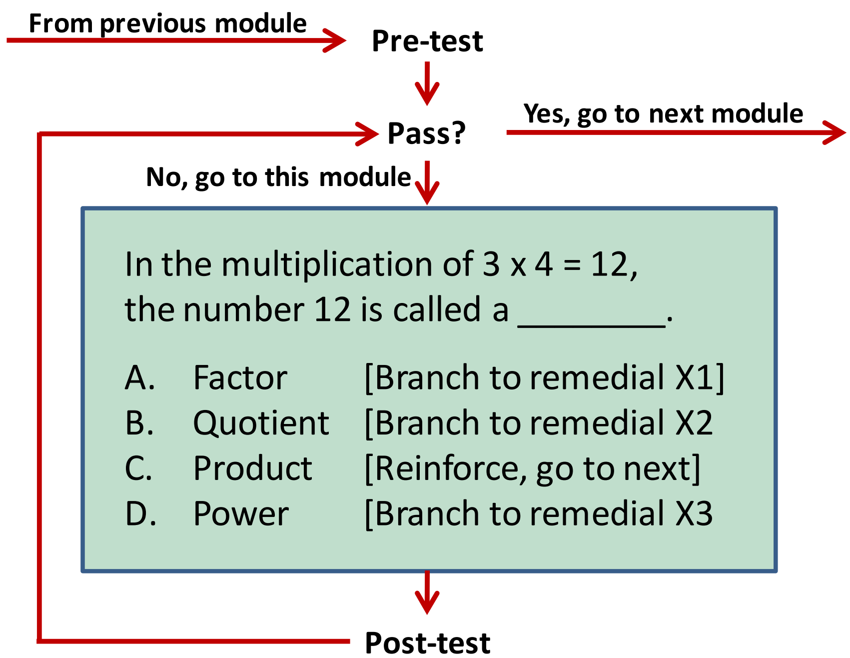

Figure 2-2: Typical Use Case – Keller's PSI Approach Combined with Crowder's Intrinsic Programming.

Reviews by Hartley [18], Kulik, Cohen, and Ebeling [19], and Kulik, Schwalb, and Kulik [20] found that this approach to CAI was more effective than classroom instruction, but only modestly so. Notably and expensively it required the developers (or 'authors') of the instruction to foresee and program for every possible state of the learner and the learning system during the course of instruction. A solution, as suggested by Feurzeig [21] and funded by US military training research [22], was for the computer to assume much of this 'authoring' and perform it in real time as needed by the state of the learner and the instructional system. Doing so, required the application of machine or 'artificial' intelligence to CAI and became the basis for the development of intelligent tutoring systems [23].

In support of this understanding in the 1960s, Feurzeig developed a programming language, Mentor, which aided instruction developers in providing this intelligence. Based on initial development of the Mentor language [21], semantic networks by Quillian [24], and his own work on SCHOLAR, Carbonell [25] identified two basic features that distinguish ITS from standard CAI:

- *Use of information structures* – Information structures are used in place of pre-programmed, frame-oriented exercises such as Crowder's [17] intrinsic programming that had been adapted from paper-based programmed learning. Information structures, such as those based on ontologies, concept maps, natural-language understanding, and one-on-one tutorial strategies, relieve developers from the need to anticipate every state that might exist for individual learners and the instructional system. Modeling these states and dealing with them were thereby assigned, as much as possible, to the computer.
- *Mixed-initiative tutorial dialogue* – Because of their generative capabilities, these systems allow either the computer-tutor or the learner to initiate relatively open-ended inquiries during an instructional dialogue. Either the student or the computer could take the initiative in asking questions, posing problems, and providing explanations. Tutorial responses by the computer could be generated and tailored to what the student had learned as well as the context within which the inputs occurred. For that matter, the computer could be prepared to provide guidance and assistance to the learner, even and if necessary, before the learner knows what questions to ask.

The following is an example extracted from computer-student interactions for training medical diagnosis skills. It was written in MENTOR programming language, as developed by Feurzeig for ITS [21], [26].

The example begins with a problem statement:

> It is mid-December. This afternoon a patient was admitted to the hospital. He is a well-developed, well-nourished, 34-year-old coal miner. He was brought in by his wife who reports that he has had "the grippe" for about 4 days. During the past 3 days she has given him fluids and aspirin every 6 hours with no sign of improvement. For the past 2 days he had complained of feeling very warm and of pain in his belly. This morning he had a severe shaking chill that was so bad his teeth rattled.
>
> <The student then asks for and receives data on the patient's vital signs and other physical exam results.>
>
> COMPUTER: Before proceeding with the reexamination, it might be instructive to pause and consider the diagnostic implications of the information at hand. Perhaps you think it too early to arrive at a definitive conclusion. On the other hand, you may be prepared to make the diagnosis now. Are you?
>
> STUDENT: No.
>
> COMPUTER: Fine. Nevertheless, we've got a fair amount of information-at least enough to raise some, perhaps several, diagnostic possibilities. Have you any in mind?
>
> STUDENT: Yes.
>
> COMPUTER: Name one.
>
> STUDENT: Pneumococcal pneumonia.
>
> COMPUTER: Good. The general signs of a pneumonic process are present. Of course, you haven't isolated the bug. Do you think the belly pain and flank pain reinforce the case for pneumococcal pneumonia?
>
> STUDENT: No.

COMPUTER: Right. Though the white count is consistent, isn't it?

STUDENT: Yes.

COMPUTER: I think so too, but a later reading is indicated to confirm this finding. We've considered one possibility. Do you have another diagnosis in mind?

STUDENT: No.

COMPUTER: In that case, I'd like to talk about pulmonary infarction. What I like about that diagnosis is the presence of high fever, rusty sputum, and shaking chills. I don't like the dyspnea, pleural pain, and signs of consolidation-do you agree?

STUDENT: No.

COMPUTER: You're quite correct. I was just trying to see whether you were alert-and you were. Now let's see how the patient is today. Begin the physical when you're ready.

<Deleted: student requests for vital signs and other physical exam items.>

STUDENT: Pneumococcal pneumonia.

COMPUTER: Yes. I remember you liked this diagnosis before. Are you absolutely sure?

STUDENT: No.

COMPUTER: What would help you nail down the diagnosis?

STUDENT: Evidence from sputum smear and culture.

COMPUTER: Good. Anything else?

STUDENT: Evidence from x-ray findings.

<The dialogue continues> [21]

This work in medical diagnosis subsequently motivated extensive work applying computers in medical education in the 1970s and 1980s by Clancey [27]; see discussion on GUIDON in Chapter 8 of this Report.

More flexible dialogue interchange in ITS systems developed during the late 1970s. It may be best illustrated by an example from SOPHIE, which was an ITS intended to teach electronic troubleshooting [28].

In this example, extracted from an authentic dialogue, the student is leading and trying to find a fault introduced into a power supply by SOPHIE.

STUDENT: What is the base emitter voltage of the voltage limiting transistor?

SOPHIE: The base emitter voltage of Q5 is -.037 volts.

<The first student-SOPHIE interaction in this example is a straightforward question and answer. SOPHIE has introduced a fault into the simulated power supply and the student is trying to locate and repair it.>

STUDENT: What about for the current limiting transistor?

SOPHIE: The base emitter voltage of Q8 is .349 volts.

<The second student-SOPHIE interaction is a little less straightforward. In addition to simulating the faulted power supply, SOPHIE must be able to parse the dialogue sufficiently to know that the student is asking for the base emitter voltage of transistor Q8.>

STUDENT: What should it be?

SOPHIE: In a working circuit, the base emitter voltage of Q8 is .593 volts.

<The third student-SOPHIE interaction continues its dependence on SOPHIE's understanding of context and dialogue. Notably the dialogue requires SOPHIE to extract answers by simulating both the faulted power supply and one that is operating correctly.>

STUDENT: Replace Q8.

SOPHIE: I am going to ask you some questions about how Q8 is faulted. Are any junctions shorted?

<The fourth student-SOPHIE interaction goes a step beyond what Brown et al. described as a knowledgeable system to what they considered an intelligent system. SOPHIE has progressed from knowledgeable parsing of dialogue with the student and simulation to an ITS with tutorial intelligence. It was important to develop an evolving, dynamic student model to manage the conversation. SOPHIE shadowed student's solution paths and prior knowledge to shadow his/her troubleshooting hypotheses. It determined that they were incorrect, elected to capture the dialogue initiative back from the student, and in this part of the dialogue begins to pursue a series of tutorial interactions to guide the student back to a correct solution.>

SOPHIE provides an example of individualization through goal-driven, tutorial, problem-solving conversations between learners and a computer-tutor. Such conversations could also point learners to pre-prepared didactic lessons, tests, simulations, and other presentations, interactions, and learning environments as needed, but whatever it did stemmed from a mixed-initiative dialogue as early envisioned by Carbonell [25] and pursued by pioneers such as John Seely Brown and Richard Burton [28]. The adaptation of instruction evident in the conversations arises from a model of the student generated by the system as the dialogue progresses and a model describing the organization of the content. Today these tutors use instructional tactics such as the following:

- Requiring logical, causal, and/or goal-oriented reasoning for both correct and incorrect responses.
- Dealing with learner impasses by reviewing the knowledge and skills individual learners have already acquired – never providing direct hints or correct answers.
- Verifying learner understanding of any didactic material before proceeding.
- Ensuring active, constant interactions to foster the immersion or 'flow' that is found in interactive computer game playing [29].
- Integrating use of human mentors to proctor learning activities and supplement computer dialogue as needed.

2.4 EFFECTIVENESS OF INTELLIGENT TUTORING SYSTEMS

Evidence of the effectiveness of ITS systems is growing. Overall they have been found to be superior to both classroom instruction and more basic CAI systems. Earlier evaluations of basic (non-ITS) CAI found that it produces about 0.33 standard deviations more learning than classroom learning over a variety of subject matter [11]. This result would be viewed as a small, but appreciable improvement, given US Department of Education guidelines.

Moreover, there are substantial variations in the findings of studies included in these reviews. Some differences in learning between CAI and classroom instruction exceed 1.00 standard deviations (roughly an improvement of 50th percentile students to the 84th percentile) and others report near zero standard deviations of difference. A review by VanLehn [7] found ITS of a certain type (step-based ITS) produced an average 0.76 standard deviations (roughly an overall improvement of 50th percentile students to about the 78th percentile). Kulik and Fletcher [30] found nearly the same result, reporting an overall improvement among 39 carefully selected ITS of 0.75 standard deviations. The findings again varied widely from slightly negative to others in excess of 2.00 standard deviations (roughly an overall improvement of 50th percentile students to about the 98th percentile). Kulik and Fletcher also noted that the poor performance of some ITSs may have been due to inadequate implementation (limited support for classroom teachers whose students used the ITS) in some cases and poor alignment of assessment tests with the objectives of the ITS being assessed in other cases.

Any newly implemented instructional capability for existing courses requires attention. The introduction of an ITS is no different. Interactive preservice instruction should be provided for using the ITS, along with daily reports on the progress of the learners toward instructional goals. Drop-in visits with instructors to answer questions, provide support, and assist in using the system should also occur while it is in use.

Table 2-1 is relevant with regard to alignment and the Anderson and Krathwohl [8] instructional space discussed at the beginning of this chapter. Kulik and Fletcher [30] report three instances where a single ITS was assessed on two separate occasions, once with assessment focused on rudimentary learning (basic facts, nomenclature, and simple procedures) and another where the assessment focused on more deep, conceptual learning. The findings in Table 2-1 suggest that ITS systems are better suited to providing deep than rudimentary learning.

Table 2-1: Effect Sizes for Three ITS Systems Assessed for Deep and Shallow Learning.

Source	Concepts (Deep)	Rudiments (Shallow)
Graesser et al. [31]	0.34	0.00
Person, Bautista, Graesser and Mathews [32]	0.30	0.03
VanLehn et al. [33]	0.95	-0.08
Average	0.62	-0.02

2.5 UNDERSTANDING THE VALUE OF INTELLIGENT TUTORING SYSTEMS

Assessments of ITSs focus on effectiveness with good reason. The ability of any instruction to produce learning is fundamental. However, effectiveness is only half an answer for military decision makers who must decide between buying more or better training instead of spare parts, jet propulsion fuel, ammunition, or other necessities. Their decisions need to be informed on the monetary and/or operational return from their decisions by data. In most cases, they routinely receive this data – except from their investments in training. People who design, develop, and provide training rarely provide data to support investments in it. Consequently, training, especially the residential, "schoolhouse" training to qualify individuals for specific occupational specialties, is viewed, formally and/or informally, as an expense rather than an investment. In the words, of an admiral responsible for providing all US Navy materiel, military decision makers need to know what a pound of training is worth.

One way to address the value in terms of return on investment in training is to compare the costs to produce a level of expertise obtained from formal, residential ('schoolhouse') training with the costs for the years of duty station experience that includes On-Job-Training (OJT) and experience to produce the same level of expertise. An analysis of this sort comparing training using an ITS with standard classroom instruction was performed by Cohn and Fletcher [34]. Their analysis was performed on data from an ITS that was developed for the US Navy by the US Defense Advanced Research Projects Agency (DARPA). After 16 weeks of training in Information systems Technology (IT), its graduates, who began as novices, outscored in IT troubleshooting and knowledge similar trainees with 35 weeks of training as well as seasoned Navy sailors averaging 9.6 years of IT experience in the Fleet [35]. All effect sizes in these comparisons favored use of the DARPA ITS and nearly all exceeded 3.0 standard deviations. However, this result is highly unusual and might be explained the cumulative effect of both improvements to instructional techniques plus improvements in the quality of instructional content.

In comparing the costs for research, development, implementation, and use of this ITS with the costs of conventional training followed by 7 years on-job-training and experience for 2000 sailors a year (the established training 'pipeline' for this occupation specialty) Cohn and Fletcher found savings from using the ITS to exceed $300M per year. While this ITS was unusual and expensive to develop, it demonstrated the potential value of ITS technology and its return on investment. Confirming results were obtained using this ITS to train military veterans by comparing return to the government for the cost of using this system [36].

Use of ITS by the military appears to promise substantially superior training at least in preparing military personnel for technical duties, improving human performance, and, more importantly, increasing the likelihood of mission success. Subsequent discussion in this report, describes and provides examples of its use and value in maintenance training, collective (team) training, STEM (Science, Technology, Engineering, and Mathematics) education and training, and medical education and training.

2.6 LEARN MORE ABOUT INTELLIGENT TUTORING SYSTEMS DESIGN AND USE

To this point in our chapter we have covered the evolution of ITS design and application in discussion of tools and methods developed in the 1960 – 1980s. This was intended to:

1) Educate the reader about differences that distinguish ITS from standard CAI; and
2) Expose the reader to the myriad of steps required to begin to make the use of ITSs practical for the masses.

Subsequent chapters in this Report cover a broader, more recent context of ITS design. Chapter 3 covers a review of ITS authoring tools. Chapter 4 touches on emerging ITS technologies. Chapters 5, 6, and 7 cover tutoring in application domains for collectives (teams), STEM (Science, Technology, Engineering and Mathematics), and medical respectively. Following up on our discussion of understanding the value of ITSs, we recommend Fletcher and Sottilare [37] for methods to assess the cost/return on investment of training and educational systems.

In addition to the papers discussed and cited in this report, we recommend two additional sources to learn more about ITSs design and use:

- *Design Recommendations for Intelligent Tutoring Systems* Book Series – Volumes 1-4 (Learner Modeling, Instructional Management, Authoring Tools, and Domain Modeling. Available online free at: https://gifttutoring.org/projects/gift/documents.
- Perez, R.S., Skinner, A. and Chatelier, P. (2016). Lessons Learned from Intelligent Tutoring System Research for Simulation. In H.F. O'Neil, E.L. Baker & R.S. Perez. (Eds.), *Using games and simulation for teaching and assessment*. Routledge: Abingdon, Oxon, UK.

2.7 REFERENCES

[1] Gettinger, M. and White, M.A. (1980). Evaluating curriculum fit with class ability. *Journal of Educational Psychology*, *72*, 338-344.

[2] Thorndike, E.L. (1906). *Principles of Teaching*. New York, NY: A.G. Seiler & Company.

[3] Tobias, S. (1989). Another look at research on the adaptation of instruction to student characteristics. *Educational Psychologist, 24*, 213-227.

[4] Tobias, S. (2003). Extending Snow's conceptions of aptitudes. *Contemporary Psychology, 48*, 277-279.

[5] Bloom, B.S. (1984). An introduction to mastery learning. In, J.H. Block (Ed.), *Schools, Society, and Mastery Learning* (pp 4-14). New York: NY Holt, Rinehart and Winston.

[6] Evans, M. and Michael, J. (2006). *One-on-one Tutoring by Humans and Computers*. Mahwah, NJ: Lawrence Erlbaum.

[7] VanLehn, K. (2011). "The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems." *Educational Psychologist 46*(4), 197-221. doi: 10.1080/00461520.2011.611369.

[8] Anderson, L.W., Krathwohl, D.R. and Bloom, B.S. (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. Columbus, OH: Allyn & Bacon.

[9] Fletcher, J.D. (1992). Individualized systems of instruction. In M.C. Alkin (Ed.), *Encyclopedia of Educational Research* (6th ed., pp. 613-620). New York, NY: Macmillan.

[10] Jamison, D.T., Suppes, P. and Wells, S. (1974). The effectiveness of alternative instructional media: A survey. *Review of Educational Research, 44*, 1-67.

[11] Kulik, J.A. (1994). "Meta-analytic studies of findings on computer-based instruction." In E.L. Baker and H.F. O'Neil, Jr. (Eds.), *Technology assessment in education and training* (pp. 9-33). Hillsdale, NJ: Erlbaum.

[12] Niemiec, R. and Walberg, H.J. (1987). Comparative effects of computer-assisted instruction: A synthesis of reviews. *Journal of Educational Computing Research, 3*, 19-37.

[13] Vinsonhaler, J.F. and Bass, R.K. (1972). A summary of ten major studies on CAI drill and practice. *Educational Technology, 12*, 29-32.

[14] Gick, M.L. and Holyoak, K.J. (1980). Analogical problem solving. *Cognitive Psychology 12*, 306-55.

[15] Koedinger, K.R., Corbett, A.T. and Perfetti, C. (2012). The Knowledge-Learning-Instruction Framework: Bridging the Science-Practice Chasm to Enhance Robust Student Learning. *Cognitive Science, 36*, 757-798.

[16] Keller, F.S. (1968). Goodbye, teacher ... . *Journal of Applied Behavior Analysis, 1*, 79-89.

[17] Crowder, N.A. (1959). Automatic teaching by means of intrinsic programming. In E. Galanter (Ed.), *Automatic Teaching: The State of the Art* (pp. 109-116). New York, NY: John Wiley & Sons.

[18] Hartley, S.S. (1977). *Meta-analysis of the effects of individually paced instruction in mathematics* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses database (Order No. 7729926).

[19] Kulik, J.A., Cohen, P.A. and Ebeling, B.J. (1980). Effectiveness of programmed instruction in higher education: A meta-analysis of findings. *Educational Evaluation and Policy Analysis, 2*, 51-64.

[20] Kulik, C-L.C., Schwalb, B.J. and Kulik, J.A. (1982). Programmed instruction in secondary education: A meta-analysis of evaluation findings. *Journal of Educational Research, 75*, 133-138.

[21] Feurzeig, W. (1969). *Computer systems for teaching complex concepts* (BBN Report 1742). Cambridge, MA: Bolt Beranek & Newman, Inc.

[22] Fletcher, J.D. and Rockway, M.R. (1986). Computer-based training in the military. In J.A. Ellis (Ed.). *Military Contributions to Instructional Technology* (pp. 171-222). New York, NY: Praeger Publishers, 1986.

[23] Fletcher, J.D. (2009). Education and training technology in the military. *Science.* 323(5910), 72-75. doi: 10.1126/science.

[24] Quillian, M.R. (1969). The teachable language comprehender: A simulation program and theory of language. *Communications of the ACM, 12*(8), 459-75.

[25] Carbonell, J.R. (1970). "AI in CAI: An artificial intelligence approach to computer assisted instruction." *IEEE Transactions on Man-Machine Systems 11*(4), 190–202. Retrieved from http://ieeexplore.ieee.org/document/4081977/.

[26] Feurzeig, W. (2011). Educational technology at BBN. In D. Walden & R. Nickerson (Eds.), *A culture of innovation*. East Sandwich, MA: Waterside Publishing.

[27] Clancey, W.J. (1987). *Knowledge-based tutoring: the Guidon Program*. Cambridge, Massachusetts: MIT Press.

[28] Brown, J.S., Burton, R.R. and DeKleer, J. (1982). Pedagogical, Natural Language and Knowledge Engineering in SOPHIE I, II, and III. In D. Sleeman and J.S. Brown (Eds.), *Intelligent Tutoring Systems*. New York, NY: Academic Press.

[29] Csikszentmihalyi, M. (1990). *The Psychology of Optimal Experience*. London: Harper Perennial.

[30] Kulik, J.A. and Fletcher, J.D. (2016). Effectiveness of Intelligent Tutoring Systems: A Meta-analytic Review. *Review of Educational Research, 86*, 42-78.

[31] Graesser, A.C., Moreno, K., Marineau, J., Adcock, A., Olney, A., Person, N. and Tutoring Research Group (2003). AutoTutor improves deep learning of computer literacy: Is it the dialog or the talking head. In U. Hoppe, F. Verdejo, and J. Kay (Eds.), *Proceedings of Artificial Intelligence in Education* (pp. 47–54). Amsterdam: IOS Press. Retrieved from http://www.cs.pitt.edu/~chopin/references/tig/AIED graesser-2003.pdf.

[32] Person, N.K., Graesser, A.C., Bautista, L., Mathews, E.C. and Tutoring Research Group. (2001). Evaluating student learning gains in two versions of AutoTutor. Artificial intelligence in education: AI-ED in the wired and wireless future, 286-293.

[33] VanLehn, K., Lynch, C., Schulze, K., Shapiro, J.A., Shelby, R., Taylor, L. and Wintersgill, M. (2005). The Andes physics tutoring system: Lessons learned. *International Journal of Artificial Intelligence in Education 15*(3), 147–204. Retrieved from http://www.andestutor.org/Pages/AndesLessonsLearnedFor Web.pdf.

[34] Cohn, J. and Fletcher, J.D. (2010). What is a Pound of Training Worth? *Proceedings of the InterService/Industry Training, Simulation and Education Annual Conference*. Arlington, VA: National Training and Simulation Association.

[35] Fletcher, J.D. and Morrison, J.E. (2014). *Accelerating Development of Expertise: A Digital Tutor for Navy Technical Training* (D-5358). Alexandria, VA: Institute for Defense Analyses.

[36] Fletcher, J.D. (2014). *Digital Tutoring in Information Systems Technology for Veterans: Data Report* (D-5336). Alexandria, VA: Institute for Defense Analyses.

[37] Fletcher, J. and Sottilare, R. (2014). Cost Analysis for Training & Educational Systems. In R. Sottilare, A. Graesser, X. Hu, and B. Goldberg (Eds.) *Design Recommendations for Intelligent Tutoring Systems: Volume 2 - Instructional Management*. Army Research Laboratory, Orlando, Florida. ISBN: 978-0-9893923-2-7.

Chapter 3 − A REVIEW OF INTELLIGENT TUTORING SYSTEM AUTHORING TOOLS AND METHODS

Robert A. Sottilare, Ph.D.

3.1 INTRODUCTION

As noted in Chapter 1, there are several barriers to the widespread adoption of Intelligent Tutoring Systems (ITSs) for military training and education. Some of the most significant challenges are in the area of authoring, the process of building and deploying ITSs for use by learners. Many authoring tools require specialized skills to use them effectively. However, authoring tools can be designed to intelligently guide the author or automate steps in the authoring process based on the skill level of the author [1], [2], [3]. The authoring of adaptive instruction delivered by ITSs requires a basis for adaptation. Adaptation, also known as tailoring or branching, is usually based on the states (e.g., performance or emotional states) or traits (e.g., personality or demographic data) of the learner. More adaptation means a higher degree of tailoring usually requires more content and more effort to author vs. a non-adaptive instructional experience (current training scenarios). Many authoring tools require the author to have knowledge and skills in three areas [4]:

- Expert knowledge of the domain to develop learning objectives and measures, and select/organize appropriate content.
- Understanding of instructional design principles to sequence lessons, manage flow, and adapt to traits and states of the learner(s).
- Software programming skills develop domain-specific interfaces (e.g., user dashboards).

How can we make ITS authoring practical for the masses? While many military training centers have staff with these skills (or can hire personnel with these skills), the efficiency of the ITS authoring process could greatly improve if the authoring tools were able to automate or incorporate some or all of these skills. From a practical standpoint, military training center staffs have expert knowledge of the domains they teach or for the instruction they generate. Our task is to examine how existing authoring tools incorporate methods for retrieving and organizing content, guide authoring to adhere to best instructional design practices, and minimize/eliminate the need for software programming.

3.2 REVIEW OF ITS AUTHORING TOOLS AND DESIRABLE CAPABILITIES

A review of ITS authoring tools follows which incorporates discussion and identification of desirable authoring capabilities. We have adapted criteria specified by Woolf and Hayes-Roth [5], [6] which included tutor authoring functions in four areas: communication, teaching knowledge, domain/student knowledge, and knowledge representation (see Authoring Functions and Subfunctions in Table 3-1). To this we added associated authoring methods and skill/knowledge required for the author to perform these functions. We examined four widely used ITS authoring toolsets with respect to their ability to support these authoring functions: The Generalized Intelligent Framework for Tutoring (GIFT; developed by the US Army Research Laboratory), AutoTutor (developed by the University of Memphis, the Authoring Software Platform for Intelligent Resources in Education (ASPIRE; developed by the University of Canterbury in New Zealand); and the Cognitive Tutor Authoring Tools (CTAT; developed by Carnegie Mellon University). The intent of this examination was not to compare these ITS tools to each other, but to provide a least common denominator to highlight critical ITS authoring functions.

Table 3-1: ITS Authoring Functions, Methods Required, and Skills/Knowledge Required.

Authoring Functions & Subfunctions	Authoring Methods Required	Authoring Skills/Knowledge Required
Communication Function		
Interface Design	allow authors to apply usability heuristics and scaffolding to support an engaging user experience; automate processes wherever possible to reduce authoring workload	ability to visually express hierarchy, grouping, and workflows for authoring
Pedagogical Agents	allow authors to specify the reactive, proactive, and cooperative behaviors of pedagogical agents; automate processes wherever possible to reduce authoring workload	ability to specify corpora and agents to activate tutor-learner dialogue; ability to identify best instructional practices and implement them as agent policies
Natural Language Dialogue	allow authors to specify conversational elements and criteria for activation of conversational agents (e.g., virtual humans)	ability to specify corpora for natural language understanding and generation including selection of the task domain and size
Instructional Knowledge Function		
Sequencing of Content	allow authors to select content to support tutoring of concepts	knowledge of the hierarchical relationships between domain concepts
Planning Interventions	allow authors to specify concepts, measures, and associated tutor interactions; assess the effectiveness of interaction and develop reinforcement learning models to continously improve effectiveness	knowledge of the domain concepts, measures, misconceptions, and options for feedback and remediation
Control Structure	allow authors to provide reasoning to: optimize tutor performance (e.g., problem solving speed and accuracy); efficently interact with learner (e.g., frequency of learner interventions)	knowledge of meta-data to describe data attributes (e.g., difficult vs. easy questions) and support tutor decisions (e.g., adaptations or branching in course flow)
Domain/Learner Knowledge Function		
Procedural Skills	allow authors to identify facts, plans, concepts, and relationships	knowledge of the domain and associated successful plans/paths and learner misconceptions
Learner Affect & Engagement	allow authors to specify methods for learner data acquisition and the development of machine learning classsifiers	ability to develop or activate classifiers based on learner behaviors or physiology; understanding of affective states and their impact on learning
Learner Misconceptions	allow authors to define misconceptions based on expert knowledge or provide executable learner models that predict learner misconceptions based on behavioral patterns	knowledge of common domain errors or misconceptions
Knowledge Representation Function		
Constraints, Rules, Grammars	allow authors to specify production rules, decision trees, instructional policies, or models for tracing	ability to construct mechanisms to guide the actions of the tutor
Scripts and Plans	allow authors to specify/develop scripts, paths, plans to identify procedural knowledge	ability to identify procedural knowledge (knowledge used in the performance of a task)
Probabilistic Reasoning	allow authors to define logic to handle uncertain situations (e.g., learner states of affect or engagement)	knowledge of uncertainty in the learner and the task domain; knowledge of methods to probabilisitically model the learner and the domain

The following subsections provide an overview of each authoring tool set, a discussion of their functions, and finally, recommendations for desirable characteristics for ITS authoring tools for NATO countries based on our criteria.

3.2.1 Generalized Intelligent Framework for Tutoring (GIFT) Authoring Tools

A unifying *GIFT Authoring Tool (GAT)*, developed by the Learning in Intelligent Tutoring Environments (LITE) Lab at the US Army Research Laboratory, currently consists of a user-friendly front end (course creator view) which link several open-source authoring tools which support the development of domain knowledge (e.g., concepts, measures, associated tutor responses, surveys, and question banks) and course flow. GIFT authoring is a data driven process with authoring data structures in Extensible Markup Language (XML). In addition to building the domain the GAT also allows for the configuration of sensors, learner models, and pedagogy. A recent evaluation of the GAT User Experience (UX) drove the recent redesign of the GAT to provide an intelligently guided authoring experience [7]. The GAT is available as part of the free GIFT downloads [8] and may be hosted as either a local instance (on your laptop) or used as a cloud-based instance (browser access). The GAT provides server-based access to the AutoTutor Script Authoring Tool, ASAT, to support dialogue-based interactions (e.g., reflective dialogue to assess learning) during tutoring.

The course file developed through the GAT describes the flow of the domain session. The flow of a course can be fixed or dynamic. A fixed course flow is simply one linear sequence without branching. A dynamic or branching course specifies one or more branching transitions (adaptations by the tutor) that use learner state characteristics and metadata to present appropriate content during the course. The course authoring view of the GAT provides a whiteboard to allow the author to sequence course flow. Standard drag-and-drop objects to represent steps in the course flow include:

- Information object − displays content (text, verbal content or a webpage) to the learner.
- Survey object − used to collect information (demographic or assessment) about the learner; the results of surveys can be used by GIFT to adapt the course flow (see adaptive course flow object below).
- Structured review object − useful for displaying survey/text responses and overall real-time assessment scores to the learner; the contents of the review are automatically collected from experiences in the course.
- Media object − delivery mechanism for multimedia content (local or web-based).
- Adaptive course flow object − separates and manages knowledge comprehension and skill into a four part experience that delivers content and provides feedback/remediation based on learner state attributes.
- External application interface object − links to previously integrated environments; allows GIFT to assess the learner’s skill in an application (e.g., Microsoft PowerPoint, serious games like Virtual Battle Space or Virtual Medic, AutoTutor Script Authoring Tool).

GIFT represents domains as concepts (objectives to be learned) and guides the learner’s experience through: techniques (e.g., error-sensitive feedback) which are instructional policies managed by software agents; strategies (e.g., engage the learner in a reflective dialogue) which are domain-independent instructional plans for future tutor action based on learner states; and tactics (e.g., launch a reflective dialogue to ascertain the learner’s understanding of concept A) which are domain-dependent instructional actions by the tutor based on strategies and constrained by techniques. All of the domain-dependent knowledge components are in GIFT’s domain module. GIFT authoring tools require no knowledge of software programming or instructional design to develop effective ITSs. GIFT is freely available and may be hosted as either a local instance or a cloud-based instance. GIFT-based tutors have been prototyped to support training in adaptive marksmanship, land navigation, medical casualty care, and other military and non-military domains. GIFT, like other ITS technologies, has been largely focused on training individual learners, but research is ongoing to create tools and methods to support the tutoring of teams. GIFT provides an interface for the commercial Media Semantics virtual character and research is ongoing to support integration and authoring in the Virtual Human Toolkit [9]. Additional information about GIFT and the GAT is available at www.GIFTtutoring.org along with free executable software code downloads for registered users.

Functional recommendations for authoring tools evolving from the examination of GIFT follow:

- Interface design – provide simple interfaces for easy integration of ITS with external environments (e.g., virtual humans, training simulations or serious games) to support the reuse of existing training and educational assets for adaptive instruction.
- Interface design – include drag-and-drop interface with standard objects (e.g., media, information or survey objects) to allow authors to quickly author course flow.
- Interface design – provide a mechanism to allow authors to copy and paste whole courses, lessons, objects, or any sub-elements of a course to reduce authoring time.
- Pedagogical agents – develop a multi-agent architecture to manage instructional policies in real-time based on best practices found in the literature (e.g., mastery, error-sensitive feedback, faded worked examples); providing this multi-agent architecture will reduce authoring workload by specifying a set of domain-independent policies which are available at run-time.
- Sequencing of content – provide a mechanism for storing, retrieving, and managing content to reduce authoring time.
- Control structure – standardize a set of meta-data labels for data attributes to support tutor decisions.
- Control structure – develop methods to rank and queue pending actions by the tutor.
- Procedural skills and learner misconceptions – investigate a reinforcement learning mechanism which can use learner data across a population to automatically identify learner processes (procedural knowledge) and misconceptions.
- Probabilistic reasoning – investigate and implement Markov Decision Processes to support probabilistic reasoning for tutorial path planning through a course or lesson.

3.2.2 AutoTutor Authoring Tools

AutoTutor has been applied to several task domains in support of one-to-tutoring, and it has a comprehensive set of authoring tools and services. The *AutoTutor Authoring Tools* are used to develop interactive tutors where students are taught through natural language discourse. AutoTutor serves as a guide that assists learners in expressing verbal content through discourse processing acts (e.g., pumps, hints, prompts). It seeks to simulate the discourse patterns and pedagogical strategies used by expert human tutors [10].

AutoTutor was developed to support specific domains (e.g., Newtonian physics and computer literacy). As the name suggests, the AutoTutor Script Authoring Tool (ASAT) is a tool within the AutoTutor framework used to create AutoTutor scripts. ASAT-X is an Extensible Markup Language (XML)-based tool. The ASAT-V tool is used to view and test AutoTutor visual scripts created by Microsoft Visio. Authoring conversation rules in ASAT can be very challenging for instructors, course managers, and domain experts. However, the AutoTutor Lite authoring interface is more intuitive and the tools are available as binary code (executable code). The major parts of an AutoTutor script include agent definition, common speech act classification, speech packs, and rules.

An emerging capability in AutoTutor is the *trialogue*, intelligent pedagogical agents that help students learn by holding a conversation in natural language between the student, a virtual instructor and a virtual student (peer) [11]. Trialogues may be authored using ASAT. Additional information about AutoTutor, its variants, and ASAT is available at www.autotutor.org/.

Functional recommendations for authoring tools evolving from the examination of AutoTutor follow:

- Interface design and pedagogical agents – develop more intuitive authoring interfaces to support natural language conversations between learners and virtual humans (e.g., dialogues and trialogues).
- Interface design – allow author to specify interfaces with other tutoring architectures (e.g., GIFT).
- Natural language dialogue – implement mixed initiative dialogue to allow either the learner or the tutor to initiate/control the conversation.
- Scripts and plans – define procedural knowledge for the task under instruction through simple scripts and plans.
- Learner affect and engagement – model and maintain the relationship between learner capabilities and task difficulty to maintain engagement and reduce unwanted affect (e.g., boredom, anxiety, frustration).
- Sequencing of content – reduce authoring workload by developing easy authoring interfaces to allow the application and sequencing of content in a task domain.
- Probabilistic reasoning – provide logic to support uncertainty in understanding natural language; reduce uncertainty by reducing the size of the domain corpus or dividing it into several logical elements.

3.2.3 ASPIRE Authoring Tools

ASPIRE is an authoring environment for developing constraint-based ITSs, which can be used by instructors to author ITSs to supplement their courses. ASPIRE supports authoring of the domain knowledge. The use of this knowledge is critical to development of the domain model which is the most complex and time-consuming part of ITS development. Constraint-based tutors model instructional domains at an abstract level, an approach that simplifies the development of ITSs by representing domain knowledge in the form of constraints, which specify what should be true, rather than generating specific problem-solving paths [12], [13].

For example, a haiku, a traditional Japanese poem, has three constraints: it always has three lines, it always has seventeen syllables, and the three lines always have 5, 7, and 5 syllables respectively. A feedback message from a constraint-based tutor like ASPIRE which is linked to these constraints communicates to a learner that their solution is incorrect, points out why it is incorrect, and reminds the learner of the corresponding declarative knowledge (i.e., the domain principle that is violated by the solution). The feedback in this case might be: '*if your third line of your poem has more or less than 5 syllables, you need to adjust it to 5 syllables to meet the definition of a haiku*'.

A goal of ASPIRE is to allow non-computer programmers to be able to author effective ITSs. Since it is understood that parts of the authoring process may be beyond the capabilities of instructors, course managers, and domain experts, ASPIRE-Author uses automation and intelligent support to provide scaffolding and guide authors through a seven step authoring process:

1) Modeling the domain structure – this step is supported by the Domain Structure Modeler which allows the author to specify the general characteristics of the selected task domain.
2) Composing an ontology of the domain – this step uses the Ontology Workspace to specify the domain ontology (formal naming and definition of the types, properties, and interrelationships of entities or things in a domain).

3) Modeling the problem and solution structures – this step uses the Problem/Solution Structure Modeler to specify the structure of problems and solutions in the domain.
4) Designing the student interface – this steps uses the Student Interface Builder to specify the initial version of the student interface which communicates (e.g., presents problems and content) with the learner relative to their conformance of the constraints.
5) Adding problems and solutions – this step uses the Problem/Solution Editor to develop and edit associated tasks (e.g., algebra problems) and their solutions.
6) Generating constraints (syntax and semantic) – all the information specified prior to this step is used by the Constraint Generator to develop the domain knowledge needed to understand the learner's solutions to the tasks/problems worked.
7) Validating the generated constraints – Finally, the generated constraints are validated in this step.

Additional information about ASPIRE is available at http://aspire.cosc.canterbury.ac.nz/.

Functional recommendations for authoring tools evolving from the examination of ASPIRE follow:

- Control structures – use constraints to minimize and simplify control mechanisms to reduce the authoring workload.
- Procedural skills – define a process for authoring including how domain knowledge is defined, structured, and accessed.
- Natural language dialogue and control structure – allow learners to stop the dialogue between them and the tutor at any point in time once they have reached a solution through their own reasoning.
- Control structure and constraints, rules, and grammars – develop an ontology, a formal naming and definition of the types, properties, and interrelationships of entities in a domain, to support the generation of constraints in that domain.
- Constraints, rules, and grammars – constraint-based tutors offer a simpler model of the domain knowledge and thereby reduce the authoring workload.
- Constraints, rules, and grammars – use constraint-based modeling to avoid the need to model learners misconceptions (errors or behaviors that indicate a lack of understanding) and thereby reduce authoring workload.

3.2.4 Cognitive Tutor Authoring Tools (CTAT)

CTAT is a suite of ITS authoring tools for developing and delivering tutors that enables you to learn by doing (i.e., active learning). The CTAT baseline is been around for many years and continues to evolve. CTAT can create ITSs to support both simple and complex problem-solving. Tutors built with CTAT provide step-level guidance for complex problem solving activities as well as individualized task selection based on a Bayesian student model [14]. CTAT tutors track learners as they work through problem sets (e.g., mathematics or physics) and then provide context-sensitive, just-in-time help.

The authoring process requires definition of a task domain along with a set of appropriate problems. CTAT was developed to support problem-based task domains. It may be more difficult to support the authoring of scenario-based tutors where problem-solving processes are less linear and multiple paths to success are the norm. In order to develop a domain model in CTAT, a cognitive task analysis is required to understand how students learn required concepts and evolve their skills. CTAT requires familiarity with the Java Expert System Shell (JESS) production rule language to develop a cognitive tutor.

Authoring tools for both cognitive and example-tracing tutors are included with CTAT. Cognitive tutors use a cognitive model to provide feedback to students as they are progress through the task of solving problems. An author can build a rule-based cognitive ITS either through Artificial-Intelligence (AI) programming [15] or by using a non-programming module called SimStudent [16]. Example tracing tutors evaluate learner behavior and compare it to generalized examples of problem-solving behavior. An author can create an example-tracing tutor using non-programming methods in CTAT [17].

CTAT is part of three part system that includes TutorShop and DataShop [18]. CTAT provides tools for authoring tutor behavior as well as run-time support for tutors. TutorShop is a web-based learning management system for ITSs and DataShop is a large online repository for educational technology data sets plus a broad suite of analysis tools, designed for use by researchers and geared towards data-driven refinement of knowledge component models underlying tutors [19].

CTAT is currently available as binary (executable) code at: http://ctat.pact.cs.cmu.edu/.

Functional recommendations for authoring tools evolving from the examination of CTAT follow:

- Interface design − seek flexibility in the ITS architecture selected to be able to create different types of tutors (e.g., cognitive, model-tracing, constraint-based or example-tracing) suitable for different domains.
- interface design − practice use-driven tool development by learning from each user's experiences, and letting user experiences be a driving factor in development.
- Interface design and pedagogical agents − seek general architectures able to create tutors covering many domains and a range of pedagogical approaches (e.g., guided invention, collaborative learning, simulation-based learning, learning from erroneous examples, and game-based learning.
- Learner misconceptions − identify existing (e.g., SimStudent) or discover new capabilities to reduce authoring workload by automating the modeling of the learner including misconceptions.

3.3 CONCLUSIONS

The authoring of ITSs is largely done as part of the development of a single capability. This makes ITSs difficult and expensive to build requiring specialized skills and complex tools. However, there are some existing authoring toolsets that are widely used which provide a diverse set of desirable capabilities for training and education in NATO countries. Some of the design considerations and desirable capabilities for ITS authoring are list in section 2 of this Chapter and should be used to drive future selection or development of ITS authoring systems for NATO use. By way of consolidation, we offer generalized recommendations for the design of ITS authoring systems in three areas: automation, user experiences, and interoperability.

- *Enhance automation* − identify/develop and use AI-based authoring capabilities to reduce authoring workload; instances where AI-based tools might be applied in ITS authoring include, but are not limited to:
    - Ontology and course flow − auto-generation of hierarchical mapping of concepts from text-based sources (e.g., field manuals).

- Domain modeling and scenario generation − auto-generation of child domains or specific scenarios based on a set of primitive features (and feature variability) in a parent domain.
- Learner modeling − unobtrusive data acquisition and classification techniques to identify learner states (e.g., learning, performance, engagement or affect) and traits (e.g., common domain misconceptions or personality).
- Authoring guidance − use AI techniques (e.g., virtual coach) to guide authors through a standard authoring process.

- *Improve user experiences (UX) during authoring* − design and evaluate authoring UX with respect to common usability heuristics so the authoring system is flexible and able to match both the mental models of users and common system capabilities; recommended UX design considerations follow:
  - Copy and paste − allow authors to copy ITS course elements at varying levels of granularity (e.g., course, lesson, module) to reduce authoring workload.
  - Drag and drop − allow authors to drag and drop standard objects (e.g., media, information, external application interfaces or sensors) to author course flow (sequencing).
  - Knowledge and content management − allow authors to store, retrieve, and organize knowledge and content as they see fit to reduce time to locate and apply information during the authoring process.
- *Improve the interoperability of ITSs* − standardizing even a few elements of ITSs in the NATO community will reduce the time, cost, and skill required to author ITSs by enhancing the frequency that existing ITSs and their models/components will be reused:
  - Interface design − standardization of gateways and message protocols to reduce the effort required to interact with existing training and educational systems in support of adaptive instruction.
  - Interoperability standards − develop a standardization effort through the Institute of Electrical and Electronics Engineers (IEEE) or International Standards Organization (ISO) with goal of instituting a Standardisation Agreement (STANAG) for ITSs.

NATO countries should begin the task of identifying desirable open source capabilities for adoption and use by the NATO community. This includes ITS authoring tools, software, models, methods, algorithms, standards, and datasets. GIFT could potentially provide these needed capabilities [20]-[21].

3.4 REFERENCES

[1] Murray, T. (1999). Authoring intelligent tutoring systems: An analysis of the state of the art. *International Journal of Artificial Intelligence in Education*, 10(1):98-129.

[2] Murray, T. (2003). An Overview of Intelligent Tutoring System Authoring Tools: Updated analysis of the state of the art. *Authoring tools for advanced technology learning environments*. 2003, 491-545.

[3] Murray T. Chapter 2: theory-based authoring tool design: considering the complexity of tasks and mental models. In: Sottilare, Graesser A, Hu X, Brawner K, editors. Design recommendations for intelligent tutoring systems. Vol. 3: authoring tools and expert modeling techniques. Orlando (FL): Army Research Laboratory (US): 2015; ISBN 978-0-9893923-7-2.

[4] Sottilare, R., Graesser, A., Hu, X. and Brawner, K. (2015). Introduction to Authoring Tools and Methods for Intelligent Tutoring Systems. In R. Sottilare, A. Graesser, X. Hu & K. Brawner (Eds.) *Design Recommendations for Intelligent Tutoring Systems: Volume 3 – Authoring Tools & Expert Modeling Techniques*. Army Research Laboratory, Orlando, Florida. ISBN: 978-0-9893923-7-2.

[5] Woolf, B.P. (2009). Chapter 10 – Future View. In *Building Intelligent Interactive Tutors: Student-centered strategies for revolutionizing e-learning*. Elsevier, Inc. Oxford, United Kingdom. ISBN: 978-0-12-373594-2.

[6] Hayes-Roth, F. (1987). Expert Systems. In S. Shapiro (Ed.) Encyclopedia of Artificial Intelligence. Wiley & Son.

[7] Ososky. S. and Sottilare, R. (2016). A Heuristic Evaluation of the Generalized Intelligent Framework for Tutoring (GIFT) Authoring Tools. *Army Research Laboratory* (ARL-SR-0353), March 2016.

[8] Sottilare, R.A., Brawner, K.W., Goldberg, B.S. and 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.

[9] Hartholt, A., Traum, D., Marsella, S.C., Shapiro, A., Stratou, G., Leuski, A. and Gratch, J. (2013, August). All together now. In *International Workshop on Intelligent Virtual Agents* (pp. 368-381). Springer Berlin Heidelberg.

[10] Graesser, A.C., Wiemer-Hastings, K., Wiemer-Hastings, P., Kreuz, R. and Tutoring Research Group (1999). AutoTutor: A simulation of a human tutor. *Journal of Cognitive Systems Research*, 1, 35-51.

[11] Feng, S., Stewart, J., Clewley, D. and Graesser, A.C. (2015, June). Emotional, Epistemic, and Neutral Feedback in AutoTutor Trialogues to Improve Reading Comprehension. In International Conference on Artificial Intelligence in Education (pp. 570-573). Springer International Publishing.

[12] Ohlsson, S. Constraint-based student modeling. Journal of Artificial Intelligence and Education, 3(4), 1992, 429-447.

[13] Mitrovic, A., Martin, B. and Suraweera, P. (2007). Intelligent tutors for all: Constraint-based modeling methodology, systems and authoring.

[14] Aleven, V., Sewall, J., Popescu, O., van Velsen, M., Demi, S. and Leber, B. Chapter 22: reflecting on twelve years of ITS authoring tools research with CTAT. In: Sottilare R, Graesser A, Hu X, Brawner K, editors. Design recommendations for intelligent tutoring systems. Vol. 3: authoring tools and expert modeling techniques. Orlando (FL): *Army Research Laboratory (US)*; 2015. ISBN 978-0-9893923-7-2.

[15] Aleven, V. (2010). Rule-Based cognitive modeling for intelligent tutoring systems. In R. Nkambou, J. Bourdeau and R. Mizoguchi (Eds.), Studies in Computational Intelligence: Vol. 308. *Advances in intelligent tutoring systems* (pp. 33-62). Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-14363-2_3.

[16] Matsuda, N., Cohen, W.W. and Koedinger, K.R. (2005). Applying programming by demonstration in an intelligent authoring tool for cognitive tutors. In D. Oblinger, T. Lau, Y. Gil and M. Bauer (Eds.),

AAAI workshop on human comprehensible machine learning (technical report WS-05-04) (pp. 1-8). Menlo Park, CA: AAAI Association.

[17] Aleven, V., Mclaren, B.M., Sewall, J. and Koedinger, K.R. (2009). A new paradigm for intelligent tutoring systems: Example-tracing tutors. *International Journal of Artificial Intelligence in Education*, 19(2), 105-154.

[18] Koedinger, K.R., Baker, R.S.J.D., Cunningham, K., Skogsholm, A., Leber, B. and Stamper, J. (2010). A data repository for the EDM community: The PSLC datashop. In S. Ventura, C. Romero, M. Pechenizkiy and R S.J.D. Baker (Eds.), *Handbook of educational data mining* (pp. 43-55). Boca Raton, FL: CRC Press.

[19] Aleven, V. and Koedinger, K.R. (2013). Knowledge component approaches to learner modeling. In R. Sottilare, A. Graesser, X. Hu and H. Holden (Eds.), *Design recommendations for adaptive intelligent tutoring systems* (Vol. I, Learner Modeling, pp. 165-182). Orlando, FL: US Army Research Laboratory.

[20] Brawner, K., Holden, H., Goldberg, B. and Sottilare, R. (2012, December). Recommendations for modern tools to author tutoring systems. In *The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC)*.

[21] Sottilare, R., Graesser, A., Hu, X. and Brawner, K. (2015). Introduction to Authoring Tools and Methods for Intelligent Tutoring Systems. In R. Sottilare, A. Graesser, X. Hu & K. Brawner (Eds.) *Design Recommendations for Intelligent Tutoring Systems: Volume 3 – Authoring Tools & Expert Modeling Techniques*. Army Research Laboratory, Orlando, Florida. ISBN: 978-0-9893923-7-2.

Chapter 4 − A REVIEW OF EMERGING INTELLIGENT TUTORING SYSTEM TECHNOLOGIES

Vladimir Zotov, Ph.D. and Thomas Alexander, Ph.D.

4.1 INTRODUCTION

Advances in Intelligent Tutoring System (ITS) development would not be possible without recent technological innovations and trends, which open new opportunities for incorporating new methods and tools into a tutoring system. One of the main engines that drives the current rise in capability and popularity of ITS applications has its roots in the proliferation of mobile, hand-held devices, which are approaching similar computational power, memory, and storage capacity to desktop personal computers. This development allowed deployment of sophisticated software packages which run on these compact devices (e.g., tablets and smartphones), thus enabling ITS applications to expand beyond their traditional use in the classrooms and computer labs to much wider fields of deployable training solutions.

These applications allow mobile learning anywhere, anytime. This is of particular interest for researchers interested in traditional knowledge transfer based on training objectives and the goal of developing individual competencies [1]-[2]. As a tool for one-to-one tutoring, ITSs facilitate active participation and engagement of the trainee which supports learning, competency development and transfer of skills [3]. The new mobile technologies allow information access without restrictions of time and/or space. Many smartphone owners use their device anytime and just about everywhere; during transportation (i.e., at airports, train stations, during public transportation) or while walking. Access to relevant training content and information is possible whenever the trainee is ready and motivated to learn.

By introducing "training nuggets", small chunks of instruction, there is potential to enhance the motivation to engage in training content by fitting short training sessions into the trainee's schedule at convenient points. Therefore, mobile technologies may increase motivation and, thus, training success. Today, mobile devices are ubiquitous. It is expected that the market share of telecommuting and corresponding devices for information access will continue to increase. Information technology developers and providers have shared this optimistic estimation of the growth potential for a decade now and it still continues. Although mobile learning is attractive and self-paced, by itself it lacks the feedback and sound pedagogic reasoning needed to provide effective, tailored instruction. ITSs on mobile technology are likely candidates to fill this role.

The shift from traditional desktop computers toward mobile devices is not restricted to tablets and smartphones. In the last few years, wearable devices have also increased in popularity. Among them are smart glasses (e.g., Google Glasses [4]) that display information within the user's field of view and communicate with the internet via natural language voice commands. Smartwatches or wrist bands provide users with important notifications. Regardless of the mobile device you have, the sensor technology in it can provide some essential information needed to allow ITSs to formulate next steps (e.g., feedback, direction or scaffolding). These mobile devices can remotely access information, track the user's global positioning, and process/storing data remotely.

Cloud computing has been introduced as a new paradigm for mobile applications, in which, instead of running mobile software on mobile devices, it is transferred to a distributed and powerful computing platform in the cloud. Centralized applications on clouds may be accessed using a wireless connection. Standardization of mobile applications under two major platforms (Android OS by Google and iOS by Apple) contributed to the rapid development of applications that are usually compatible cross-platform with increasingly high assortment of handheld and wearable devices.

In parallel, cyber security has become an important issue for mobile information access. This is crucial for educational content that is accessible from both restricted and public services. A constraint in the use of mobile devices for tutoring is likely the inability to allow classified information on wireless and cloud-based training networks.

While the portable computers and mobile devices brought the required computation power that enabled ITS to run complex and challenging software, another technological boost comes from the advances in sensor development. The cost of varies sensors from heart rate monitoring to wearable EEG fell drastically and they are becoming the mainstream consumer gadgets. Miniaturized traditional sensors for capturing physiological data, (including heard rate and galvanic skin response to name just a few) found their ways to the latest generation of wearable devices equipped with multiple sensors that allow users to track physical activities, distance, calories burned, sleep patterns, and even stress level (e.g., Muse [5]). Sensors are important part of ITS as they allow to collect objective data of human states and input it back into the ITS to adjust and fine-tune the training feedback. These new neuro-physiological adaptive ITSs provide aids to the learner's progress and skill level based on physiological measures and biological markers of learning. The feedback can improve training process and boost trainees' performance.

When a tutoring application is for training a skill that includes physical responses it is usually designed to be used in a synthetic training environment, such as flight simulator or a submarine maintenance training Virtual Reality (VR). Virtual Reality is a special kind of simulation. NATO HFM-021 [6] has defined Virtual Reality (VR) as:

> "*... the experience of being in a synthetic environment and the perceiving and interacting through sensors and effectors, actively and passively, with it and the objects in it, as if they were real. Virtual Reality technology allows the user to perceive and experience sensory contact and interact dynamically with such contact in any or all modalities*."

VR-technology has already become widely available in civil applications and for diverse military applications. In addition to pure technology and hardware, VR also depends on the comprehensiveness and fidelity of underlying digital models. The requirements for fidelity are very high in terms of visual and acoustic presentation of the environment. Today, VR is being implemented and used successfully at different military training facilities. With serious gaming and commercials-of-the-shelf it is possible to facilitate efficient and effective training, and make it available at different or even distributed locations. The effectiveness of these applications will not only be dependent on the extent to which AR/VR meets the ergonomic requirements of human operators but also on the individual experiences with similar technologies. This will also result in new requirements for training and in changes for learning management [7].

Not only simulators have been advancing, they became less expensive and more compact. Their use is very attractive for military due to their cost, compactness, and applicability in situation when a skill involves dangerous work. A closely related technology is augmented reality where virtual reality is overlaid with real world. Despite the fact that use of VR and augmentation was tested and tried in early ITS models (e.g., Ref. [8]), it was not widely used due to such technological limitations as high cost, low display resolution, and slow visual update, making the entire experience of using VR cumbersome while triggering adverse physiological reactions known as VR sickness when wearing head mounted displays [9]. The latest VR devices (e.g., Oculus Rift [10], Play Station VR [11]) are inexpensive, high-resolution, fast-refreshing devices that diminish (but not completely eliminate) adverse physiological responses.

Incorporating elements of computer games into educational process (known as "serious gaming") is not a new trend and has been used in educational field for relatively long time (for review, see Ref. [12]). Recently, the rapid development of technologies described above along with a gradual demographic shift among trainees − who are often experienced gamers [13] − brought new trends where training and educational applications are installed on hand-held devices allowing students to learn and practice training materials any time and place they wish to.

Up until recently, most ITSs have been built through extensive knowledge engineering, and ideally with aim of cognitive task analysis, which are necessary to develop models of student and expert skill and performance. These models rely on knowledge representations, "production rules" or "constraints" that require extensive programming, expertise, and often empirical research to develop. Good examples of these models are SOAR and ACT-R [14], [15]. While the development of new approaches in AI are not as noticeable and profound as technological advances, there is also shift from traditional, rule-based AI toward statistical models based on Bayesian principles (e.g., Ref. [16]). New data-driven methods can enable more rapid development of new intelligent tutoring systems that rely on statistical, data-driven, machine learning approaches for automated or semi-automated development of the key components and functionalities of ITS.

The follow-up section provides examples of the recent intelligent tutors that incorporate some of the new trends in technology and science. Considering limitations of the scope of the report, it is not possible to provide a comprehensive review of all major intelligent tutors that are currently available. Instead, a small number of applications developed for military and educational use will be presented. These systems are good representatives of the recent technological trends encompassing many of their features.

4.2 EXAMPLES

The following examples are descriptive of prototype systems which may or may not have transitioned for use in the training of operational forces. They were selected to illustrate the application of ITS technologies in a variety of task domains. The emerging ITS technologies discussed herein are intended to be illustrative and not exhaustive.

4.2.1 Immersion and Practice of Arousal Control Training (ImPACT)

Teaching stress management skills to soldiers in Canadian Arm Forces (CAF) was an objective of a project responsible for creating Immersion and Practice of Arousal Control Training (ImPACT) [17]. The ImPACT aimed to overcome the limitations of the traditional approaches to Stress Management Training (SMT), which lack of practice of the skills in stressful situations and of soldiers' resistance with approaches related to emotions. The ImPACT tutor uses VR and video games to provide a solution to these problems. The tutoring system relies on physiological signs of stress obtained by Heart Rate (HR) and Skin Conductivity (SC) sensors, which allows the system to evaluate the trainee's level of stress. Figure 4-1 displays a diagram of the major component of the ImPACT system. The system was implemented in a Cave Automatic Virtual Environment, providing a highly-immersive, 180-degree visual environment enhanced by VR glasses for 3D experience. Additional investigations revealed that although there are plenty of available options to implement a training system in an existing VG, the openness of the VG dictates what is feasible. The ImPACT is not a typical training system with clearly defined ITS components. Nevertheless, the learner's model is represented by the tutor-user interface. The state of learner is defined by her stress level and performance. This state is compared to a state of an "ideal" player (e.g., calm, high-performing) and the discrepancy is addressed by the tutor's prompts to initiate tactical breathing and apply some other stress-reducing techniques. Additionally, the difficulty level of game is adjusted and at some graphical enhancements inform player regarding her stress level. The system allows certain degree of stress-tolerance depending on the experience of the learner; the state of the system can be adjusted by changing ImPACT parameters. The ImPACT interface is represented by configuration file allowing some basic authoring.

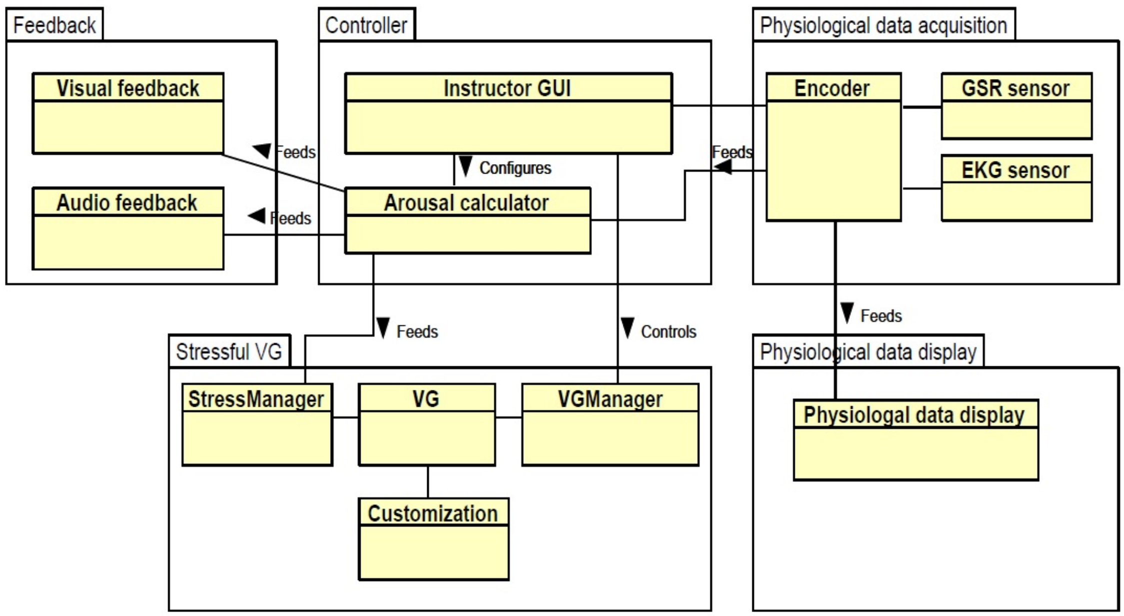

Figure 4-1: ImPACT Diagram.

An experiment tested whether ImPACT would increase the efficacy of SMT compared to usual training. ImPACT training consisted in three daily sessions to practice SMT while using biofeedback to provide information on current level of arousal and while being immersed in a VG to induce stress. Statistical analyses of the HR and SC revealed significantly less stress among soldiers who have trained with ImPACT. In addition, these soldiers performed better and were more confident in their stress coping skills. As a result, they are more likely to apply tactical breathing in the future. Finally, a survey confirmed the attractiveness of ImPACT for users. The immersion in a stressful VG, coupled with biofeedback, was shown to be more effective than training as usual and it has the additional advantage of favouring the "buy-in" of emotion management tools by soldiers.

4.2.2 Moving Target ITS

Moving Target ITS (MT-ITS) is a good example of adopting ITS as an embedded training tool in virtual environment [18]. The application was developed at Defence Research and Development Canada (DRDC) to test whether new suite of weapon-mounted tools can improve marksmanship performance of CAF infantry personnel. To evaluate the effectiveness of the FCS as a training tool, shooting performance for Moving Targets (MTs) was assessed with a portable, computer-based training program. The purpose of this program was to coach MT shooting and to act as a testbed for investigating the effectiveness of the Fire Control System (FCS) as-a-training-tool concept. This coaching system was embedded in the weapon sight of the training simulator. It was based on the FCS's correct point of aim calculation feature and on AI that delivers training tips and feedback, similar to the way a human tutor or coach would. The correct point of aim provided by the FCS helped teach trainees on adjusting their target lead according to the distance and speed of the MT, whereas the ITS system provided feedback based on participant performance, showing errors and providing tips to correct errors.

The MT-ITS was implemented in the Virtual Immersive Soldier Simulator (VISS) platform running the Virtual Battle Space 2 (VBS2) simulation software. The VISS is a first-person shooter test bed in which the participant stands approximately 10 feet from a set of three 10-foot screens on to which the environment generated by VBS2 is projected from the participant's perspective. The shooter is equipped with a mock rifle (an airsoft M4) mounted with markers that are being tracked by an optical tracking system. This system tracks where the weapon is pointing in the virtual environment, providing a link between the virtual environment and the physical weapon. Because the simulator knows where the weapon is pointing in the virtual world, the image corresponding to the sight's field of view of the virtual world can be displayed in the weapon sight in real time, with appropriate magnification. The displayed image was part of the simulated environment with magnification. VBS2 is a military first-person-shooter simulation based on a popular gaming platform. The underlying program has been modified so that key dependent measures (e.g., completion times, number of shots fired, type of target engaged, fire accuracy) can be logged for later analysis. VBS2 simulates the visual and auditory aspects of a tactical operational environment. While the aim of MT-ITS to become a stand-alone, fully embedded system, it is still in early stages of development and it relies on the VISS environment, in which it is implemented. The learner model is derived from the learner's shooting performance: efficiency of fire (number of rounds per hit), accuracy of fire (hit/miss ratio), centre of mass, response time, etc. These values are compared with the reference values, which is part of the expert model (obtained from SMEs in CAF Infantry and CAF marksmanship instructors). MT-ITS evaluates the discrepancy between learner and expert models and, depending on the type of deficiency, a customized feedback is generated to address the learner deficiencies and the ways to overcome them. At this point, MT-ITS only supplies the learner with detailed feedback and tips; the follow-up version will customize training exercise on-the-fly in response to specific deficiency.

The experiment was designed to test the effectiveness of MT-ITS by comparing shooting performance before and after MT shooting training. Three types of marksmanship training were evaluated: a control group received no training between pre- and post-training sessions, a standard CAF training group received typical MT training as described in the CAF's C7-C8 handbook, and an advanced training group received MT-ITS training between sessions. Given the military, and specifically the dismounted infantry, nature of the task, all participants were from Canadian Army units (both Regular Force and Reserve) and most were qualified with the C7A2: thirty-six had The Personal Weapons Test 3 (PWT3) or higher training. For the MT-ITS training condition, participants were shown animated hostile entity moving across the range (Figure 4-2, bottom panel). On every other trial, participants were shown the correct POA in the form of a small red square joined to the centre of mass of the target by a red line. Feedback, in the form of shot location, hits/misses, POA, and general tips reflecting participants' performance and advising them on how to improve their shooting accuracy was provided after each trial of MT-ITS condition. This condition represents what could be achieved by an embedded training system. Overall, participants' performance in MT-ITS condition was superior to two comparative groups: a control group that received to training between pre- and post-sessions, but also to a group that received a standard CAF marksmanship training.

Figure 4-2: A Replication of the Standard CAF Target (Top Panel) and Animated “Hostile” Target (Bottom Panel) as Seen Through the Digital Sight.

4.2.3 QuestionIT

QuestionIT is another training tool that was developed at DRDC to improve handling of improvised explosive device disposal [19]. The tutor has a built in instructional content and Situational Awareness (SA) clues into questions and questioning strategy to efficiently reveal situational clues as a critical means of Improvised Explosive Device (IED) detection. QuestionIT provides interactive feedback based on the student response. This intelligent attribute depends on built-in scripts (i.e., instructional content and SA clues in the questions and questioning strategies). The QuestionIT necessitates a multi-modal psychophysiological metric assessment for determining at minimum attention and engagement of the trainee using Electroencephalography (EEG) and eye tracking. Heart Rate Variability and Electrodermal Activity as indirect measures of workload have been recommended for the ITS. EEG data correlated with training content events to develop workload measures provided the information needed to classify student cognitive learning styles and determine the appropriate low-resolution sensors. Figure 4-3 shows the main components of the ITS. As the Figure shows, QuestionIT is driven by interaction between Learner and Expert models, which is guided by the system's Adoption engine with inputs from the learner's state.

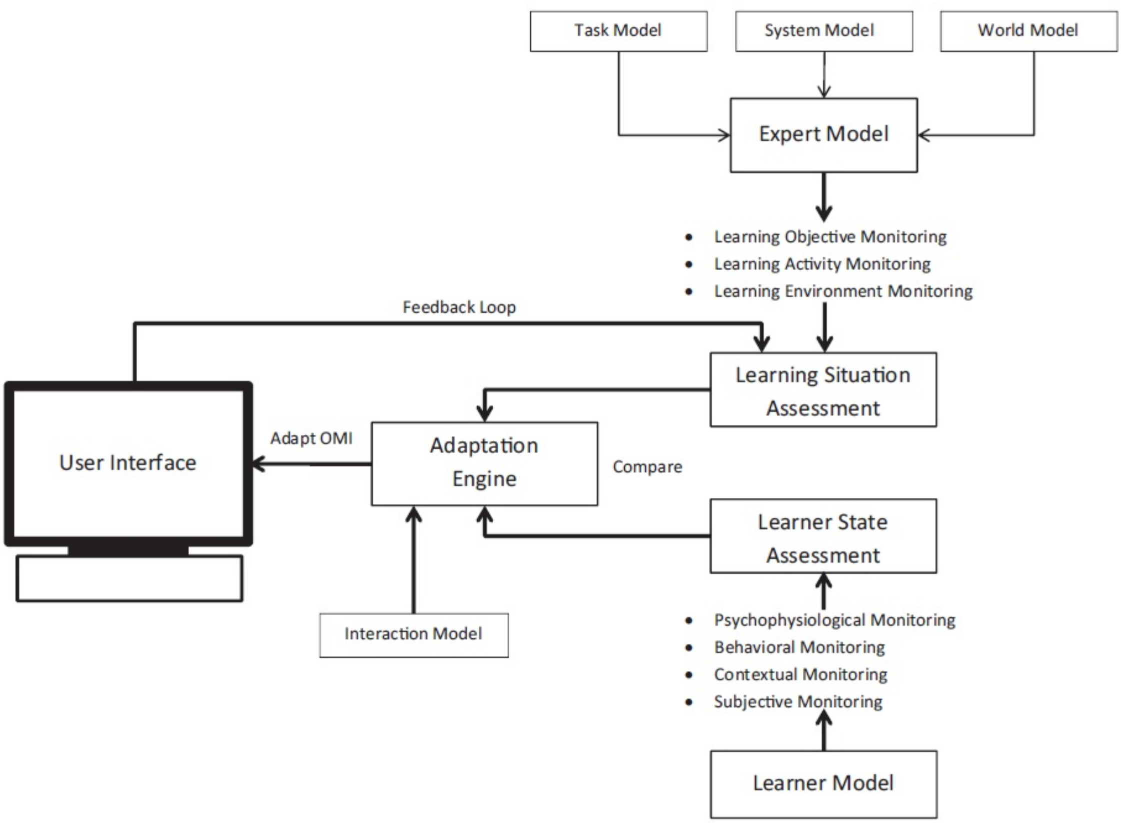

Figure 4-3: QuestionIT Components.

Comparison between a no-tutor version of QuestionIT and tutor-enabled version served to evaluate the training efficacy of the adaptive feedback and instruction capabilities. One of the main performance evaluations is questioning efficiency which is determined by a ratio of the number of SA clues revealed to the number of questions asked. Other measures included realism, usefulness and ease of use of QuestionIT. The results showed that trainees using the tutor enabled version asked fewer, yet relevant questions with improved questioning efficiency. A success rate for a following course after QuestionIT was developed and used in IEDD training school was improved to 94%. The overall results do suggest that the trainees' became better with the tutor-enabled QuestionIT and this positive gain transferred into the live role-play aspect of the training.

4.2.4 Virtual Leo 2

Virtual Leo 2 is an example of team-oriented, human-in-the-loop ITS that was designed for training of Leopard 2 crew commanders within a Virtual Leopard project at DRDC [20]. It was built on the principles of complex virtual operators using the Simulated Operator for Networks architecture (SimON) [21]. SimON is a deterministic, rule-based AI agent model, which was an appropriate type of model to store and retrieve domain knowledge in a well-structured operational environment of the Leopard 2 crew.

The Virtual Leo 2 contains a multi-agent model for the tank crew: a gunner, a driver and a loader model. By combining the operator behaviour in the Leopard 2 simulation videos and domain knowledge implemented in the Task Network Models (TNMs) using Unity3D placeholder, the tutor allows verbal exchanges and visual representation of action for the tank crew. Virtual Leopard 2 relies on speech production and recognition software. This functionality allows simulating simple Leopard 2 training drills, helping Leopard 2 crew commander to practice drills with virtual crew. The Hierarchical Task Analysis (HTA) method was used to develop the crew task network model, and the Agile software development methodology was employed for the software design and development. The knowledge obtained by HTA represents the structured domain of the basic tasks that Virtual Leo 2 was designed to address. The evaluation of the crew commander verbal input generates learner model, which is evaluated and compared with expected command as defined by HTA. The discrepancies are sorted and assigned to specific AI crew agent, which then generate a verbal reply to the crew commander. This verbal reply might be a confirmation and acceptance of the command or it can be a request for clarification if the command is not disambiguate or not compliant with HTA. While still in testing stages, the Virtual Leo 2 tutor showed how ITS can be adopted for team-based training, providing both training and experimentation testbed.

4.2.5 Dragoon

Dragoon is a basic intelligent tutoring system developed by a team of researchers at School of Computing, Informatics, and Decision Science Engineering of Arizona State University. It was designed to teach the design and development of models of dynamic systems [22]. The modelling skill is important part of K12 curriculum of math and science students. Dragoon is a so-called "step-based" tutoring system, which relies on two nested loops in its architecture:

a) Outer loop that executes *once-per-task* and maintains an assessment of the student and recommends a follow-up task (e.g., viewing videos, reading texts or answering a multiple-choice question) for the student to maximize the student's learning and motivation.

b) Inner loop that executes once-per-step and might give immediate or delayed feedback and hints on that step. Dragoon also relies on the mechanism of example-tracing: for each solved problem the system provides an example of the solution, which is based on comparison of the student's steps to steps in the author's solution.

The model has author and student modes, which serve to evaluate student solution, compare it with the author's solution and, if necessary, provide corrective feedback. The decision whether to provide student with corrective feedback depends on the model's pedagogical policy, which is a table that pairs conditions with responses.

The tutor was evaluated in a number of trials, where K12 students were tested on their ability to construct complex models. Students worked in groups of two or three and used either Dragoon or paper workbooks. The results varied from no statistical difference between groups to clear advantage of Dragoon ($d = 1.02$, $d = 0.62$, and $d = 1.00$). As the authors note, Dragoon is still work-in-progress finished, but it does show how a relatively simple tutoring system can be very effective teaching students complex tasks. For detailed information about Dragoon architecture, documentation, and source code visit GitHub website at http://github.com/bvds/LaitsV3.

4.2.6 BEETLE II

The Basic Electricity and Electronics Learning Environment (BEETLE II), developed under the direction of the US Naval Warfare Center's Training Systems Division and jointly with the University of Edinburgh, has advanced computer-based tutorial dialogue capabilities to support conceptual learning in electricity and electronics. BEETLE II is capable of interpreting student natural language explanations to support active experimentation, self-explanation, and generation principles. This tutor provides relevant one-to-one tutoring in a military domain, and opportunities to reuse its tutoring processes are considered significant. Additional information about BEETLE II is available at: http://groups.inf.ed.ac.uk/beetle/beetle_bib.html.

4.2.7 Tele-Maintenance

The main military application of a research and technology project of Fraunhofer FKIE is maintenance and, thus, supporting military technical personnel maintaining and/or repairing complex technical structures. The main focus of the original project was to implement a tele-cooperation system for support under operational restrictions. However, it is currently investigated how to extend the system as an assisting system that can be applied for educating and training maintenance personnel. Linking the system to an ITS is promising for a enhancing effectiveness and efficiency. But this is ongoing research. The general importance of fast and qualitative maintenance has also grown within the armed forces because of the diversity and variety of platforms and technical equipment. This requires a broad knowledge of platforms and general technology rather than special knowledge in detail. Instead, special knowledge and details about platforms are presented either by an expert by reach-back or an assisting system. This first approach is understood by the term "tele-maintenance" [23] so that an expert cooperates with a logistics soldier. Sharing a common visual space in such a tele-maintenance session also improves communication and interactive cooperation [24]. However, the results of applying simple video communications and web-conferences are worse than a direct cooperation [25]: Error rates and time to perform a task increase. Interactive Electronic Technical Documentation (IETD) already provides detailed knowledge about complex systems, but it is often not available appropriately. By adding videos of lessons learned or 3D CAD data it is possible to enhance usability and comprehensiveness of the content, but still it is not sufficient for a thorough understanding of most complex military systems [26]. This asks for support of an ITS to transfer knowledge appropriately.

In addition to IETDs and electronic manuals it is important to connect the system with additional functionality, e.g., to ask for interactive assistance from an expert. Figure 4-4 shows an illustration of the system at Fraunhofer FKIE.

Figure 4-4: Illustration of Tele-Maintenance Implementation at Fraunhofer-FKIE [27].

The interactive cooperation usually requires sufficient bandwidth for a synchronous transfer of audio-visual information within the network. Delays or latency will reduce effectiveness and efficiency of the repair. A pure reduction to acoustics and, thus, verbal description, is no solution because it reduces performance significantly. A solution is to share the visual space by using video in a web-conference [26]. But this requires bandwidth that is not available under operational conditions. Our approach to bridge the gap is based on a virtual reconstruction of the maintenance object through the use of Augmented and Virtual Reality (AR/VR). Figure 4-5 shows the general tasks required for gaining a common understanding of the problem and developing a solution.

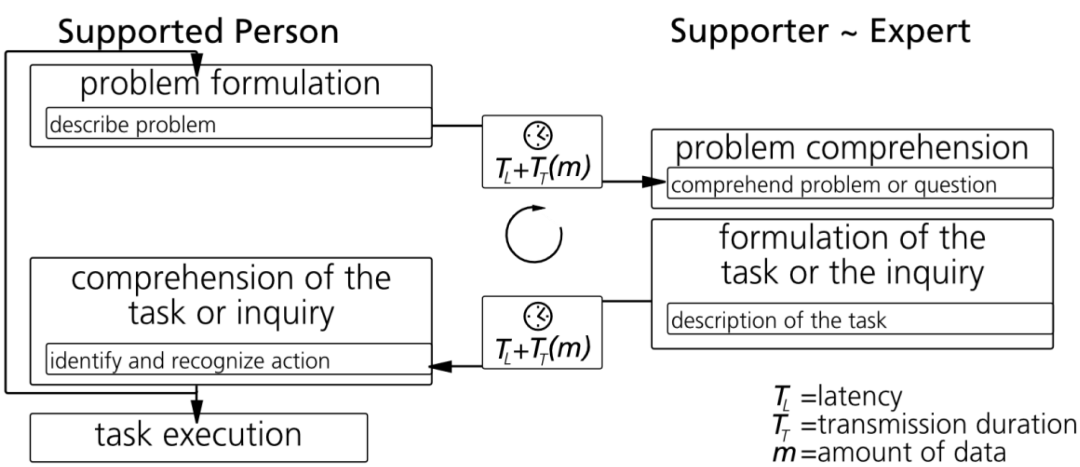

Figure 4-5: System Diagram of a Typical Maintenance and Repair Setting [27].

The results of our experiments showed that performance increased significantly. A subjective evaluation proved higher ratings for the AR-VR system than other technical solutions. We found no negative effect on subjective workload or on visual fatigue. The results therefore support our concept of a stereoscopic 3D system as a promising tool for creating instructions in a tele-maintenance task. Up to now the system was only evaluated for maintenance tasks. Yet, the concept is also applicable in the diagnostic phase. Selecting, placing and using the right diagnostic procedure and tools can be supported by a remote expert. Another area well suited for the use of an integrated AR-VR system is training or advanced distant learning. The instructor may use the VR system to instruct multiple students but can also give individual instructions [27]-[28].

The ongoing extension of the system changes the actual application from maintenance to training. In this case, the ITS takes the role of the expert (i.e., the right side of Figure 4-5). The ITS will have to identify the state of the trainee and of the learning content automatically. As a consequence it formulates tasks to perform or inquiries for the student. This requires a thorough investigation of relevant tasks and functions to be performed. However, this work is in an early stage but it uses the same hardware as the active co-operation between the two participants.

4.2.8 Tutor-Expert System Model for ITS Authoring

As we mentioned in the Introduction, while the majority of ITS packages represent knowledge using production rules, some newer system use distributed networks for domain knowledge representation. A good example of such tutor system is Tutor-Expert System (TEx-Sys) ITS [29], which was designed at University of Zagreb to teach undergraduate courses in math and science. The knowledge is represented through semantic networks with frames and production rules. Nodes in the semantic network express knowledge on subject matter objects (i.e., facts and terms) in the chosen knowledge base. Links express the process of thinking about relations among nodes of the base. The student progress evaluated and assigned to a new node (which reflects a student knowledge level) by Examination module. TEx-Sys has been used in the educational process at the Faculty of Natural Sciences, Mathematics and Education in Split, where a number of knowledge bases for different undergraduate courses have been developed.

4.2.9 Generalized Intelligent Framework for Tutoring (GIFT) for ITS Authoring

The Generalized Intelligent Framework for Tutoring (GIFT), developed by the Learning in Intelligent Tutoring Environments (LITE) Lab at the US Army Research Laboratory, is emerging as a multi-domain, open source tutoring architecture [30]. As such, it is not a specific-purpose ITS but a research prototype intended to reduce the computer skills and cost required to author ITSs, deploy them, manage them, and continuously evaluate the adaptive instruction they provide. A major advantage of GIFT is that three of its four functional elements are reusable across task domains. GIFT may also be linked to external training environments (e.g., serious games or virtual and augmented reality simulations) through a standardized gateway. GIFT authoring tools require no knowledge of computer programming or instructional design to develop effective ITSs. GIFT is freely available and may be hosted either locally or cloud-based. GIFT-based tutors have been prototyped to support training in adaptive marksmanship, land navigation, medical casualty care, and other military and non-military domains. GIFT, like other ITS technologies, has focused on training individuals, but research is underway to create tools and methods to support tutoring of collectives. At the time of this writing, GIFT has a community of over 800 government, academic, and industry users in 53 countries. Additional information about GIFT is available at www.GIFTtutoring.org.

4.3 RECOMMENDATIONS

New ITSs are becoming more efficient, effective, and diverse, showing good training results, in some instances approaching human instructors in their effectiveness [31]. The examples of successful ITS application are promising and are undoubtedly becoming valuable tools for defence purposes in a wide variety of settings, whether it is in learning a new skill, training to operate a new equipment, or mastering crew-based communication.

Nevertheless, these new technologies and applications are still in early stages of development with some limitations to overcome. Some of these limitations are technical in nature. For example, in sensor development field one of the most often cited limitations is noisy recording and inaccurate signal reading and interpretation. Even EEG, which is the most mature and advanced physiological sensor available today, is still prone to external noise interference and loss of signal reliability over time. Some less prominent limitations, which are in the process of being resolved at the time of writing this review, are related to lack of standards in evaluation of ITS training effectiveness and in their ability to provide lasting training effect due to limited number of experiments that evaluate transfer of training effectiveness. (For review, see Ref. [20].)

4.4 REFERENCES

[1] Alexander, T., Westhoven, M. and Conradi, J. (2016). Virtual Environments for competency-oriented education and training.

[2] Marstaller, A. Ausbilden für die Streitkräfte, Proceedings of the Ausbildungskongress der Bundeswehr, Hamburg: HSUniBw HH 09/2015.

[3] Bloom, B.S. (1984). An introduction to mastery learning. In, J. H. Block (Ed.), *Schools, Society, and Mastery Learning* (pp 4-14). New York: NY Holt, Rinehart and Winston.

[4] Google Glasses (2013). Retrieved from Google Glasses Website: https://www.google.com/glass/start.

[5] Muse (2016). Muse, the Brain Sensing Headband. Website: http://www.choosemuse.com/.

[6] Alexander, T. and Goldberg, S. Virtual Environments for Intuitive Human-System Interaction - National Research Activities in Augmented, Mixed and Virtual Environments. RTO-TR-HFM-121-Part-I. NATO RTO, Neuilly-sur-Seine, FR (2005).

[7] Alexander, T., Sottilare, R. and Goldberg, S. (2014). Improving Human Effectiveness Through Embedded Virtual Simulation. Technical Report HFM-165. Neuilly-sur-Seine. NATO Science and Technology Organization.

[8] Rickel, J. and Johnson, W.L. (1997, August). Intelligent tutoring in virtual reality: A preliminary report. In Proceedings of the eighth World Conference on AI in Education (pp. 294-301). Amsterdam: IOS Press.

[9] Cobb, S.V., Nichols, S., Ramsey, A. and Wilson, J.R. (1999). Virtual reality-induced symptoms and effects (VRISE). Presence, 8(2), 169-186.

[10] Oculus VR LLC. (2015, May 20). Oculus Rift. Retrieved from Oculus Rift Website: https://www.oculus.com/.

[11] PlayStation VR (2016). Retrieved from PlayStation Website: https://www.playstation.com/en-ca/explore/playstation-vr/.

[12] Squire, K. (2003). Video games in education. Int. J. Intell. Games & Simulation, 2(1), 49-62.

[13] Connolly, T.M., Boyle, E.A., MacArthur, E., Hainey, T. and Boyle, J.M. (2012). A systematic literature review of empirical evidence on computer games and serious games. Computers & Education, 59(2), 661-686.

[14] Laird, J.E., Newell, A. and Rosenbloom, P.S. (1987). Soar: An architecture for general intelligence. *Artificial intelligence*, 33(1), 1-64.

[15] Anderson, J.R. (1990). The adaptive character of thought. Psychology Press.

[16] Koedinger, K.R., Brunskill, E., Baker, R.S., McLaughlin, E.A. and Stamper, J. (2013). New potentials for data-driven intelligent tutoring system development and optimization. AI Magazine, 34(3), 27-41.

[17] Bouchard, S., Bernier, F., Boivin, É., Morin, B. and Robillard, G. (2012). Using biofeedback while immersed in a stressful videogame increases the effectiveness of stress management skills in soldiers. PloS one, 7(4), e36169.

[18] Zotov, V. Future Small Arms Research (FSAR) WBE 1.3.3: Embedded Training of Shooting Moving Targets. L16-0620-1201. July 2016.

[19] Hou, M. and Fidopiastis, C. (2016). A generic framework of intelligent adaptive learning systems: from learning effectiveness to training transfer. Theoretical Issues in Ergonomics Science, 1-17.

[20] Guo, J. (2014). Development of Leopard Tank Crew Model. DRDC Contract C15-1209-1209. CAE.

[21] Cain, B., Magee, L. and Courtney Kersten, C. (2012). Validation of virtual environments incorporating virtual operators for procedural learning. Technical Memorandum DRDC Toronto TM 2011-132.

[22] VanLehn, K., Chaudhari, P., Desai, A., Feng, J., Grover, S., Joiner, R. and van de Sande, B. (2016). The design and development of the dragoon intelligent tutoring system for model construction: lessons learned.

[23] Sanchez, C., Fernandez, F., Vena, L.S., Carpio, J. and Castro, M. (2011). "Industrial telemaintenance: Remote management experience from subway to industrial electronics," Industrial Electronics, IEEE Transactions on, vol. 58, no. 3, pp. 1044-1051.

[24] Fussell, S.R., Kraut, R.E. and Siegel, J. (2000). "Coordination of communication: Effects of shared visual context on collaborative work", In: Proceedings of the ACM conference on Computer Supported Cooperative Work, ACM, pp. 21-30.

[25] Alexander, T., Pfendler, C., Thun, J. and Kleiber, M. (2012). "The influence of the modality of telecooperation on performance and workload," Work: A Journal of Prevention, Assessment and Rehabilitation, vol. 41, no. 0, pp. 3476-3483.

[26] S1000D 4.0:2008-08-01 (2008), "International specification for technical publications utilizing a common source database", S1000D steering committee.

[27] Kleiber, M., Winkelholz, C., Alexander, T., Flemisch, F. and Schlick, C.M. (2013). Interacting and Cooperating Beyond Space: Tele-maintenance within a Virtual Visual Space. In: Proceedings of the NATO STO Symposium on Beyond Time and Space in Orlando, FL (14.-16.10.13). Neuilly-sur-Seine: NATO CSO.

[28] Kleiber, M. and Alexander, T. (2011). Evaluation of a mobile AR tele-maintenance system. Stephanidis (Ed): Universal Access in Human-Computer Interaction. Springer Lecture Series in Computer Science. ISBN: 978-3-642-21671-8.

[29] Stankov, S., Glavinić, V. and Rosić, M. (2000, January). On knowledge representation in an intelligent tutoring system. In Proc. 4th IEEE International Conference on Intelligent Engineering Systems (pp. 17-19).

[30] Sottilare, R.A., Brawner, K.W., Goldberg, B.S. and Holden, H.K. (2012). The generalized intelligent framework for tutoring (GIFT). Orlando, FL: US Army Research Laboratory–Human Research & Engineering Directorate (ARL-HRED).

[31] VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. Educational Psychologist, 46(4), 197-221.

Chapter 5 − A REVIEW OF INTELLIGENT TUTORING SYSTEMS FOR COLLECTIVE (TEAM) TRAINING

Robert A. Sottilare, Ph.D., Paula J. Durlach, Ph.D. and J.D. Fletcher, Ph.D.

5.1 INTRODUCTION

As noted in Chapter 1 of this Report, Intelligent Tutoring Systems (ITSs) have been found to have effect sizes commonly ranging from 0.70 to 0.80 σ when compared to traditional classroom training. This is approximately the equivalent of one letter grade in most school systems. These results are similar to those produced by highly competent one-to-one human tutors and provide evidence of the potential of ITSs as effective instructional tools. ITSs primarily, but not solely, support the development of knowledge and skills in well-defined cognitive task domains such as computer programming, mathematics, and physics [1]-[2]. Until recently, ITS development has focused almost exclusively on individual learning during one-to-one tutoring experiences. Its emphasis has been on the mastery of skills in individual task domains versus the tutoring of teamwork skills that might be applied broadly across domains.

Teams may be described as groups consisting of two or more individuals who must interact with one another in order to accomplish a common task, objective, or mission [3]. Team members usually have complementary skills. A team differs from a group in that completing the objective involves interdependencies across individual members. Members may have specialized task-relevant knowledge and roles, and may have access to different information sources and/or use different tools. The roles and responsibilities of individual team members may be specifically assigned, or they may arise spontaneously, depending on team size, team leadership, and the presence of new team members [4]. These assignments include requirements for communication and coordinated action depending on the task.

Teamwork, the set of interrelated behaviors and actions that occur among team members while performing a task [5], differs in the quantity and quality of communication and coordination required. For example, a team study by Jones [6] compared baseball, tennis, football, and basketball teams by defining the effectiveness of individual team members through the lens of team effectiveness and success. The study found success to be positively associated with the effectiveness of individual members of baseball, tennis, and football teams, but not basketball teams, where success depends on more closely balanced communication, timing, and coordination among members than the other three. The greater the need for these functions, the greater the need to deal with the team as a learning unit – as a learner with its own team mental model – and a consequent greater need to develop and assess shared mental models.

Teams, the basic building blocks of most military organizations, are critical to meeting the challenges associated with military missions. While teams are important in all organizations, there is greater emphasis on teams and teamwork in military organizations. To meet the needs of teams and enhance the performance of their various missions, researchers have developed concepts to provide training at the point-of-need wherever Soldiers, Sailors, Marines, and Aviators go (e.g., mobile learning, embedded training).

Embedded training is the delivery of built-in instruction to individuals or teams onboard military platforms (e.g., ground, naval or aviation platforms). As concepts for embedded training have evolved, researchers began to envision mechanisms for guided crew training onboard military platforms [7]-[8]. For example, Germany's

Puma Infantry Fighting Vehicle embedded training system provided a distributed architecture where vehicles and live instructors were geographically separated, but instructors could provide verbal guidance to the crew based on live video and communication feeds from the vehicle [9], [10]. As embedded training concepts developed, researchers began to explore the potential of artificially intelligent agents to guide embedded team learning [11]-[13]. This led to investigations of what it would take to support a generalized concept for computer-based tutoring of teams [14]-[16] and its potential return on investment.

Recently, military trainers [17]-[18] have begun to push for intelligent instructional capabilities that support the training and education of teams as well as individual learners in a variety of task domains representing the complexity of multiple situations, conditions, and scenarios found in the operational environment. This chapter discusses team training in the military, emerging approaches to using ITSs as collective or team tutoring tools, major challenges to realizing a team tutoring capability in NATO countries, and finally, recommendations for research investments to accelerate team tutoring capabilities.

5.2 TEAM TRAINING IN THE MILITARY AND NASA

Militaries have teams and teams of teams that often operate under pressure with limited time to deal with rapidly evolving or ambiguous situations. Yet, historically, dedicated learning objectives related to teamwork skills have not been included in military training. According to Baker [19], not till the 1990s did the U.S. military put any serious resources towards research on teamwork and team training interventions.

This research led to the development of Team Dimensional Training (TDT). TDT has been employed by the U.S. Navy in multiple teams and has been adapted for use by other military and non-military organizations. TDT involves observation and analysis of a set of authentic training exercises relevant to the team and provides feedback and coaching on teamwork skills. The aim is for the teamwork learning to transfer to future team activities and to endow the team with the ability to assess and diagnose their own performance in the future. A facilitator or leader provides a pre-mission brief to clarify the teamwork training objectives (focusing on information exchange, communication, supporting behavior, and initiative/leadership). The facilitator or leader then assists the team in performing a mission while also diagnosing their performance. This is followed by structured debrief – an After Action Review (AAR) in which the team discusses their own performance, and sets goals for improvement. The team is reminded of these goals prior to each subsequent exercise. TDT and other forms of team skills training have proven to be effective [20]-[23] and may be especially useful for coalition teams consisting of personnel from different organizations and/or cultures.

Relatively less research has been conducted on applying these types of methods to collective training involving teams of teams. Yammarino *et al.* [24] describe collectives as teams of teams, units, networks, and/or organizations, where the members are interdependent based on a hierarchical structuring or a set of shared expectations and goals. They point out that understanding (and developing training for) leadership skills must not ignore the influences that higher echelons have on leaders at lower echelons. If they do, consideration of contextual conditions, which affect team leadership, will fail to be represented in models, constructs or training scenarios. For teams of teams with hierarchical relationships, teamwork across teams at the same level of the hierarchy and teamwork up and down the hierarchy both require consideration.

Caldwell [25] describes complex multi-team dynamics at the Mission Control Center at the Johnson Space Center. He describes the social, informational, and technical factors influencing team-based knowledge sharing and task coordination required for mission operations involving interdependent teams of skilled experts with non-overlapping knowledge and task roles, and requiring a high degree of synchronization while operating under strong contextual and situational constraints. Not only were astronauts thoroughly drilled in simulations prior to a mission, the Mission Control Center teams supervising missions to the International Space Station were trained this way too. However, Caldwell focuses on the complex information bottlenecks and challenges, and says little about training simulation development or assessment.

Neither within-team nor collective skills training have been widely adopted by military organizations, with the exception of Crew Resource Management (CRM) training for aviation teams [19]. One reason for their omission may be the additional time and resources required; refer to NASA's challenges in Space Flight Resource Management (SFRM) [26]. The first step is to "*understand a team's temporal rhythms and episodes and then to consider what, when, and how teamwork processes contribute to critical performance outcomes by conducting a team-level task analysis embedded within a temporal framework*" [27]. From this, an expert model of mission execution should be developed, against which team performance is compared and diagnosed [28]. Optimally, "event-based" scenarios with embedded team and task competency triggers are designed in order to ensure the exercises include opportunities to manifest the competencies to be trained. Finally methods to measure multiple dimensions of performance, both automated and manual must be developed, and human observers trained to use the manual methods [28]. In contrast to militaries, NASA has adopted team skills training as a routine part of astronaut mission training [29]. One reason for the difference in adoption may be NASA's smaller training audience and narrow tolerances for breakdowns in communication, coordination, and other key team skills in space operations. While NASA's team skills training has primarily focused on team skills most related to space mission performance (e.g., communication, coordination), with the prospects of long duration missions (such as to Mars). NASA is also looking to develop training for some of the "softer", more widely applicable, aspects of team skills, such as conflict resolution [29].

Since the development of TDT, the U.S. Navy has invested resources to make its application less resource intensive, but these have all been research efforts that have not yet been transitioned to routine use. One of these projects was the Debriefing Distributed Simulation-Based Exercises (DDSBE) Program, which was intended to address the need to automate some of the processes required in conducting team training in distributed environments [30]-[32]. The U. S. Navy's current Adaptive Training for Combat Information Centers (ATCIC) initiative continues this effort by supporting authoring of event-based scenarios and reducing instructor workload [33]. For the domain of detect-to-engage tactical decision making, the instructor is given semi-automated support for observing, assessing, and remediating TDT. The detect-to-engage chain is the process by which a team detects a potential threat, evaluates the threat, and determines the response. A general framework was developed to represent the stages of the detect-to-engage process, which could subsequently be operationalized for different specific domains. The four taskwork pillars of this framework are: Detect, Identify, Elaborate, and Manage, while the five team process pillars are Information exchange, Completeness of reports, Utilizing information from all sources, Error correction, and Backup behavior. Behavioral measures that could be easily automated (e.g., accuracy of typing in a code) were automated, while measures requiring more judgment were left to the instructor to interpret.

For any specific scenario application, authoring involves specifying scenario events intended to elicit targeted task or team-work responses and the conditions under which the events would be triggered. It also includes the ability to author alerts to the instructor indicating likely opportunities to make certain observer-based assessments (guided by checklists). An option allows a scenario event to be triggered automatically, or instead to alert the instructor who then has the option of triggering that event or not, or triggering a different event. This allows the instructor to adapt the scenario flow according the dynamic context. Scenario authoring also permits the inclusion of automated hints or prompts (delivered by a simulated commander) for anticipated lapses or errors. Thus while TDT is still challenging, the ATCIC initiative is aiming to reduce the burden on training development and evaluation.

Lessons learned from military team training should pave the way for more seamless transition to adaptive team training as challenges are addressed and team tutoring capabilities move from state-of-art to state-of-practice. The next section of this chapter highlights the few areas in which research is focused on solving some of the challenges to realizing team tutoring capabilities for NATO.

5.3 MODELING TEAM LEARNING AND PERFORMANCE

The scope and character of models that team members share differ with team objectives, the extent of the teamwork required, and roles that team-members play. At some level, however, all team-members and their models must share a common understanding of team processes, interactions, and objectives. The extent to which they do and whether or not it matters can be assessed by team success in performing tasks, objectives, and missions.

Training for teams must adapt to or even prepare for the self-organization and self-assembly that occur in all teams [4]. This preparation seems especially important for the pick-up teams that are frequently and inevitably assembled to perform military operations. Such teams initially lack the 'transactive memory' developed by members of established teams. This memory contains the knowledge and skills of specific team members and an awareness of who can perform team tasks under what conditions of motivation and support [34]. It allows for a division of cognitive labor within a team, permitting the team's collective knowledge to exceed that of any individual team member.

Studies reviewed by Moreland [35] and Lewis & Herndon [36] found a strong positive relationship between transactive memory and team performance. One reason for this finding may be the long noted inverse relationship between frequency of communication and the quality of team performance in reviews of collective behavior [37]-[38]. Communications can be minimized only if the members of teams share a common understanding of the situation and what can be done by whom. The remainder of this chapter will briefly review the team training and collaborative learning literature, explore the challenges for designing Intelligent Tutoring Systems (ITSs) for teams, and bring to light some existing examples where ITS technologies (tools and methods) have been applied to team-based task domains (e.g., collective training or collaborative problem solving).

What is the difference between a team and just a group? A team differs from a group in that a team is assembled to complete an objective, such that success involves interdependencies across individuals. Members may have specialized task-relevant knowledge and roles, and may have access to different information sources and/or use different tools.

Interdependencies involve information – the extent to which one member needs information that someone else has. They also can involve performance – the extent to which the ability of one member to complete their task depends on another's performance. Thirdly, interdependencies can be process-related—how information is shared or how tasks are coordinated. Process interdependencies may be determined at the level of the team itself, or determined at organizational levels above them. Marks, Mathieu, and Zaccaro [27] define team processes as: *members' interdependent acts that convert inputs to outcomes through cognitive, verbal, and behavioral activities directed toward organizing task work to achieve collective goals*. In summary, to be effective, team members not only need to be competent at their own tasks, they also need good communication and coordination and other teamwork skills. Teamwork skills have been characterized in different ways, but can be summarized by what Salas calls the "Seven C's" [39]:

1) Cooperation − driven by attitudes and beliefs about the team and the value of teamwork.

2) Coordination − Behavioral mechanisms including mutual performance monitoring, back-up and support, adaptability, and task-related assertiveness.

3) Communication − Information exchange protocols including clarity, timeliness, confirmation, and openness.

4) Cognition − Shared understanding of roles and responsibilities, knowledge of the objectives, knowledge of other team members (transactive memory), and awareness of how to react in different situations (cue-strategy associations).

5) Coaching − Leadership activities promoting teamwork, caring about team members, and sharing.

6) Conflict Resolution − Interpersonal skills, mutual trust, and psychological safety enabling people to be assertive about and take accountability for problems.

7) Conditions (Context, Composition, Culture) − Team norms are clear and appropriate and the team works in a supportive context. Also includes factors such as organizational and/or environmental context.

While these competencies can be taught to an individual using knowledge delivery and case studies, this chapter focuses on training at the team level. Toward this end, Landon, Vessey and Barrett [40] outline six team training strategies:

1) Event-Based Training / Scenario-Based Training − teams learn and practice team skills in a simulated operational context.

2) Self-Correction Training / Guided Self-Correction Training − teams review past performance, self-diagnose, and create plans for improving.

3) Cross-Training—teams are trained on the tasks of different team members (in order to foster a shared understanding of how each contributes to the team objectives).

4) Stress Training − Teams are taught to recognize the symptoms of stress, and to practice relaxation and other stress-reducing methods.

5) Team Adaptation and Coordination Training − a case-study method and team guided discussion.

6) Team Building − team activities (not necessarily naturalistic to the team) meant to build trust and cohesion.

If team skills are learned within an operational context in which they will be used, the learning is most likely to transfer to actual operational performance. Experts consider combining methods above, with the event-based and self-correction training being essential, to get the most benefit out of team skills training [26]. That is, teams should practice team skills in a simulated operational context, review their own performance in that context, learn to self-diagnose, and create plans for improving. This approach can be supplemented with other methods like cross-training or stress training. Therefore, an ITS for team training should be designed to operate in a dynamic, simulation environment and to facilitate team self-reflection.

In NASA's Space Flight Resource Management (SFRM) program, for example, new flight controllers receive lessons on the fundamentals of team work skills, tailored with examples and discussion specific to their particular role (e.g., flight controllers). They may then do some "part-task" simulations, and then a more involved simulation, with the opportunity for self-reflection and guided debriefing. Thereafter opportunities to apply the concepts taught during the dedicated SFRM classes are embedded (and assessed) in technical lessons and simulations [26].

Now that we have introduced the topic of team tutoring, discussed the importance of teamwork in military training, and reviewed some of the literature surrounding team learning and performance, we can examine emerging approaches to the complex problem of collective tutoring.

5.4 EMERGING APPROACHES TO COLLECTIVE TUTORING

This section examines emerging approaches to collective tutoring. This does not include concurrent tutoring of multiple individuals in individual task domains, but does include adaptive tutoring of teams by ITSs and computer-based guidance of teams engaged in collaborative learning. While the teamwork literature has received sufficient attention [41], much less is known about core team behaviors, attitudes and cognition, and their influences on team outcomes (e.g., learning and performance). Although teamwork, team performance, and team learning to have received considerable coverage in the literature, much less attention has been paid to ITSs for team training or the use of Artificial Intelligence (AI) in the education of teams (e.g., collaborative learning). Recently, however, there has been a small spike in collaborative learning and collaborative problem solving publications [42]-[44]. Given the limited concentration of research on collective tutoring, we focus on two significant efforts: adaptive tutoring of teams and adaptive collaborative learning.

5.4.1 An Emerging Approach to Adaptive Team Tutoring

Recently, the US Army Research Laboratory sponsored a cooperative review of team performance, teamwork, and ITS literature to identify antecedents to four team outcomes: learning, performance, satisfaction, and viability [16], [45]. The goal of this research was to develop a practical process for computer-based adaptive tutoring of teams. The core contribution of this research was identification of behavioral markers associated with antecedents of team performance and learning to support development and refinement of team models in ITS architectures. For ITSs to optimally tailor team instruction, ITSs must have key insights about both the team and the learners on that team.

To aid the modeling of teams, we examined the relationship of team behaviors (e.g., communication, cooperation, coordination, cognition, leadership/coaching, and conflict) to team outcomes (learning, performance, satisfaction, and viability) as part of a large-scale meta-analysis of ITS, team training, and team performance. While ITSs have been used infrequently to instruct teams, the goal of this meta-analysis was to:

a) Identify significant relationships among team behaviors, effective performance, and learning outcomes.

b) Develop instructional guidelines for team tutoring based on these relationships.

c) Integrate these team tutoring guidelines with the Generalized Intelligent Framework for Tutoring (GIFT), an open source architecture for authoring, delivering, managing, and evaluating adaptive instructional tools and methods [46].

During this work, a process for the adaptive instruction of teams began to emerge in the form of structural equation models (e.g., Figure 5-1) which empirically supported a portion of the ontology identified earlier in this chapter [14], [47].

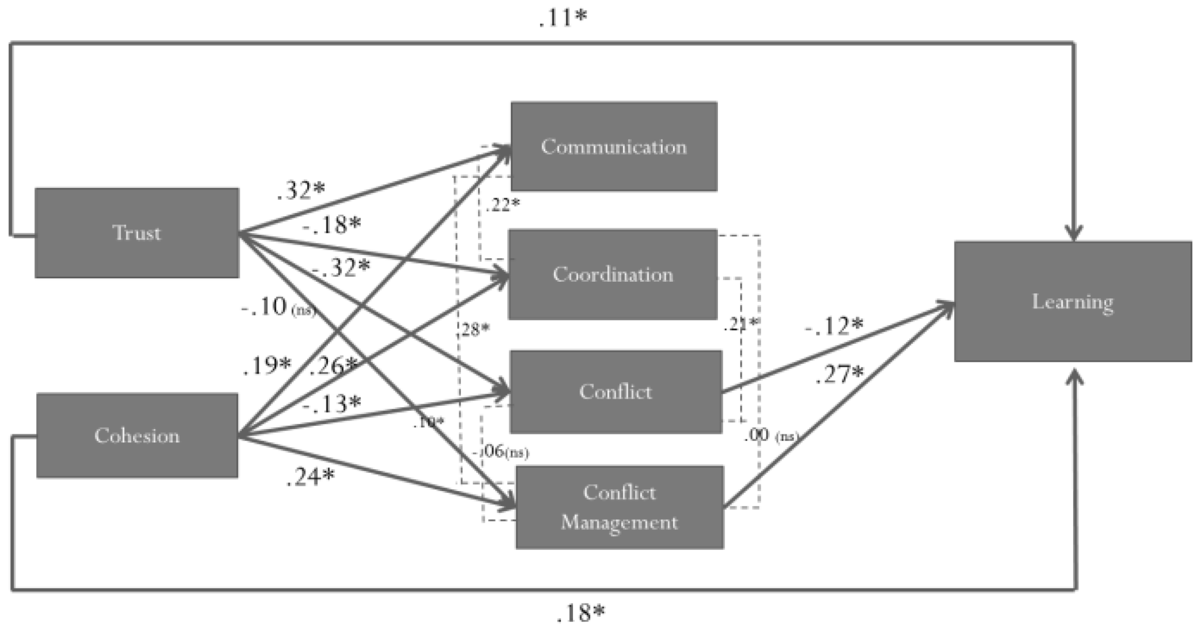

Figure 5-1: Ontology for Team Learning Based on Meta-Analytic Structural Equation Modeling [45].

An important step in this adaptive team tutoring process is the identification of behavioral markers to allow machine recognition of specific team states. While these markers were identified [45], the methods to allow machine recognition of these behaviors are still in progress. Since teamwork and team interaction is largely based in communications, it is likely that a semantic understanding of the team's verbal exchanges may be required for the next step in the tutoring process, the selection of appropriate instructional strategies [40] and tactics. In addition to team communications, much can be learned about team states from proximity behaviors and frequency of team member interactions [48], [49]. The GIFT project team expects to make strides toward this goal over the next 3 – 5 years in support of the US Army future Synthetic Training Environment (STE). Potential is high for using these tools and methods in future NATO collective ITSs.

5.4.2 Emerging Approaches to Adaptive Collaborative Learning

Before discussing emerging approaches to adaptive collaborative learning, it is useful to compare and contrast collective training and collaborative learning. The US Training & Doctrine Command defines *collective training* as "training, in either institutions or units that prepares cohesive teams and units to accomplish their combined arms and service missions throughout full-spectrum operations" [17]. A key characteristic of collective training is that a group (crews, teams, squads, and platoons) learns and executes tasks that require all elements of the group to work interdependently to meet the performance standards of the tasks. Collective training tasks are practiced under a specified set of conditions that varying in complexity (e.g., density of opposing forces, temperature, ambient light, urban vs. non-urban environments).

Dillenbourg [50] notes a broad definition of *collaborative learning* as "a situation in which two or more people learn or attempt to learn something together". He finds this definition unsatisfactory due to the many ways elements (i.e., "two or more", "learn something", "together") within the definition can be interpreted. Dillenbourg's dissatisfaction can be our flexibility. "Two or more people" can be interpreted as a pair (2), a small group (3 – 5), a class (20 – 30), a community (100 – 10,000), or a society (100,000 – millions), and any intermediate levels in between these categories. "Learning something" can allow us to broadly interpret our learning goals, and "together" can denote a variety of methods in which we share mental models within teams (e.g., face-to-face, computer-enabled, or synchronously/asynchronously).

Both collaborative learning and collective training concepts involve groups working together to learn by doing something (e.g., complete a task, solve a problem, make a decision, reach an objective, understand principles, or create some artifact), but the goals are different. *Collaborative learning* tends toward an educational approach (broadly applied knowledge; learning principles and theory) versus collective training which takes a training approach (learning to do a specific task resulting in a specific skill). Computer-mediated collaborative learning experiences may be useful for learning principles in a domain. The mediator in this case could be a virtual human interacting with the group in a shared environment (e.g., virtual world) to promote good teamwork behaviors or a human instructor supplemented by a virtual human and an instructional dashboard to allow the human instructor to monitor teamwork behaviors.

As noted earlier, there has been a recent spike in collaborative learning publications within the AI in education community. Much of this effort has been in support of collaborative problem solving goals established by the Program for International Student Assessment (PISA), an international assessment group that measures the reading, mathematics, and science literacy of 15 year old students every three years. A few approaches to adaptive collaborative learning are gaining attention in academia. Even individual learning experiences are taking on the facade of collaborative learning in the implementation of a tutoring concept called *triads* [51]. Triads are dialogue-based tutoring experiences where a human student interacts with a virtual instructor and a virtual student. This collaborative learning experience mimics what we might see in a fully human team experiencing collaborative learning.

Drury, Kay, and Losberg [52] studied learning during group computer science course assignments to gain useful insight into student perceptions about group-work over time. This effort focused on analysis of verbal interaction characteristics through data mining techniques, thereby avoiding the need for comprehensive natural language understanding capabilities within the tutor. An important assessment capability for collaborative team learning will be the ability to distill learner interaction down to easily recognizable patterns or visualizations to allow either human or computer-based instructors to classify team states and intervene as needed [53].

5.5 MAJOR CHALLENGES TO REALIZING TEAM TUTORING CAPABILITIES

Many of the preparatory activities recommended for team training are the same as those required for the development of ITS for individual training; but, that is hardly surprising, as they are also the elements of good instructional design for scenario-based training. The question is, what are the additional challenges for developing a team-based ITS, over and above those that already exist for developing an ITS for scenario-based training? This question was explored by Sottilare *et al*. [14], who discussed the challenges for designing an ITS for fostering both taskwork and teamwork in a distributed simulation environment. Some of these challenges remain the same. The following is a discussion of where we are in overcoming the challenges identified in 2011 [14] and what new challenges have been identified since.

5.5.1 Low Cost Passive Sensing of Team Behaviors and Physiological Measures

The acquisition of team behaviors (e.g., interactions between team members) and the physiological measures of individual team members has largely been overcome for desktop tutoring environments (e.g., teams operating in serious games). The commercial advent of low cost sensors like Microsoft Kinect, which can track gestures, posture, and facial markers for individuals in seated or standing positions. It is now possible to construct similar sensors through open source software using a high resolution web camera. Physiological measures (e.g., breathing and heart rates) can easily be captured and transmitted to local networks from commercial breathing straps with heart sensors. Interactions between the team members (communications and gestures) in the serious game can be tracked to determine the frequency of communication and proximity measures for their avatars.

Beyond shared virtual environments like serious games and virtual training simulators, the problem of team sensing becomes more complex. If the goal is to use ITSs for teams in live environments, the ability to track and maintain identification of individual team members is critical for accurately classifying their states and providing appropriate feedback to the right team member in near real-time. This might also be possible in instrumented spaces where individual are identified by specific color markers or other unique markers (e.g., RFID tags). Moving forward, the ability to support passive sensing in ever more complex training environments (e.g., team tutoring in un-instrumented areas also known as "tutoring in the wild") is desirable and presents new challenges (e.g., connectivity and filtering/processing of physiological data).

5.5.2 Classification of Team and Individual Team Member States

Once the team interaction data and individual learner data is acquired, the ability to filter and process the data to determine team states [39] is the next challenge. The recent identification of team behavioral markers for various states [45] is a first step in enabling the classification of those states by an ITS. An ontology is needed to break those states into more measurable primitive behaviors that can be interpreted by an ITS. It is not enough to identify a marker for team cohesion "contributing to discussions about new course of action/problem solving", it must be unambiguously measurable within the domain being trained. Assessment of the new markers will include determination of their broad applicability to training domains. For example, the occurrence of positive statements like "we make a great team" is measurable in any domain.

The classification of individual team member emotions is critical to understanding team interactions – in the context of a specific goal (e.g., team member emotions like anger and frustration are indicative of conflict within the team and chances of reaching team objectives are diminished when conflict is unresolved). While the advent of new machine learning methods have made the classification of individual emotional states possible, these classifications must be accurate and completed in near real-time. Physiological-based models are not reusable since even in the same individual physiological measures change over time due to daily changes in consumption (e.g., amount of stimulant intake life caffeine) and sleep patterns. New challenges in classification "in the wild" are related to the ability of sensors to acquire needed learner data for machine learning classification. New commercial technologies like smart glasses and smartphones offer improved opportunities to train in un-instrumented augmented reality environments where tutor-learner interaction can support adaptive instruction in land navigation [54], [55], triage and hemorrhage control [56], and marksmanship [57]. Additional research is needed to expand ITS capabilities to team task domains.

5.5.3 Selection of Optimal Instructional Strategies

Given the complex set of conditions that describe the states of learners and the environment in which they train, it is difficult to imagine how an ITS will be able to select an optimal strategy to keep team learning and performance on track. However, the state of reinforcement machine learning techniques continues to improve and the ability of ITSs to make better and better decisions also improves with it. This is now the area of most significant challenge to overcome on the road to collective tutoring. Our ability to simulate teams and their variability has the potential to accelerate progress in optimally selecting instructional strategies and tactics for any given team and set of conditions.

5.5.4 Tracking Multi-Dimensional States

As noted in the previous subsection on optimal strategies, there are several conditions representing learner and environmental states that are critical to ITS decision making. Determining which of these states has the most effect on learning and performance will help us simplify the decision process moving forward, but this is now less of a concern than it was five years ago. Experimentation is producing findings that help identify the variables of interest and their effect on team learning and performance. Challenges remain in identifying team satisfaction and viability, and more research is needed to identify their antecedents with reliable statistical power.

5.5.5 Real-Time Interaction

Any framework for a team tutor must account for optimal solutions to: where (locally or centrally via a server) states of the team are determined; how frequently they need to be reassessed; and what information is to be held locally or distributed. ITS management of this process is essential to maintaining the ability to provide timely feedback and decision making. Sensor data at least, may need to be stored and processed locally to maintain near real-time assessment and tutor-learner interaction. As GIFT and other frameworks mature, analysis of data distribution in the architecture will continue to be essential to managing workload and system performance for real-time operation.

5.5.6 A New Challenge for Collective Tutoring

What was not discussed in Sottilare *et al*. [14] is the challenge of authoring an ITS for collective tutoring. The complexity introduced by increased numbers of team members, their dynamically varying roles, and the adaptive tutoring paradigm itself increase the workload for developing collective tutors. To reduce this workload, we must seek and develop new methods to automate as much of the authoring process as possible, and we must improve the usability of our authoring tools so domain experts without software programming or instructional design knowledge can produce effective tutors. We should also be evaluating methods to scaffold or support novice ITS authors in the development of relatively simple tutors while enabling more experienced users to produce more complex tutors. As part of the authoring task, we need to provide flexible methods (e.g., widgets and dashboards) so authors can organize their domain knowledge (e.g., media, surveys).

5.6 RECOMMENDATIONS

This section recommends research and development investments needed for NATO STO to develop or adopt emerging adaptive tutoring capability for teams. Four primary recommendations follow:

- Conduct a study to identify specific ROIs for team training domains that provide near term opportunities) for applying ITS technologies to team training.
- Conduct research to discover methods to accurately classify critical team states.
- Conduct research to discover methods to select optimal team instructional strategies (ITS plans for action) and tactics (ITS actions).
- Identify options for adoption as adaptive collective tutoring standards (e.g., authoring, message sets, domain-independent models) to enhance interoperability and reuse within and between NATO countries.

5.7 CONCLUSIONS

The Return On Investment (ROI) for ITSs that guide individual learning is well documented. A primary component of this ROI is the ability of the ITS to accelerate individual learning during adaptive instruction. Based on this, we expect to see similar returns for ITSs that guide the learning of teams during adaptive instruction. To realize these returns, several challenges must be overcome. Primary among challenges are the ability to unobtrusively acquire team data and individual learner data to accurately classify team states. Meeting these challenges will provide opportunities to understand and model teamwork behaviors and then effectively use ITSs to guide teams during adaptive instruction.

We imply that ROI in individual tutoring is multiplied many times in team contexts given that militaries generally operate in team contexts and that team performance can be much more than the sum of individual performances. Finally, we note that opportunities to accelerate team learning and/or reduce time lost due to training inefficiencies will result in cost avoidances, increased throughput, and reduced training infrastructure costs. We recommend that NATO nations attempting to implement ITSs for teams understand the ROI for various individual and collective tasks before embarking on an ITS development program.

5.8 REFERENCES

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

[2] Kulik, J.A. and Fletcher, J.D. (2016). Effectiveness of Intelligent Tutoring Systems: A Meta-analytic Review. *Review of Educational Research, 86,* 42-78.

[3] Salas, E. and Cannon-Bowers, J.A. (2000). The anatomy of team training. *Training and retraining: A handbook for business, industry, government, and the military*, 312-335.

[4] Guimera, R., Uzzi, B., Spiro, J. and Amaral, L.A.N. (2005). Team assembly mechanisms determine collaboration network structure and team performance. *Science*, 308(5722), 697-702.

[5] Salas, E., Burke, C.S. and Cannon-Bowers, J.A. (2000). Teamwork: emerging principles. International *Journal of Management Reviews,* 2(4), 339-356.

[6] Jones, M. (1974). Regressing group on individual effectiveness. *Organizational Behavior and Human Performance*, 11, 426-451.

[7] Zachary, W., Cannon-Bowers, J.A., Burns, J., Bilazarian, P. and Krecker, D. (1998). An advanced embedded training system (AETS) for tactical team training. Paper presented at the 4th International Conference on Intelligent Tutoring Systems, San Antonio, TX.

[8] Burmester, G.M., Stottler, D. and Hart, J.L. (2005). Embedded training intelligent tutoring systems (ITS) for the future combat systems (FCS) command and control (C2) vehicle. Army Research Development and Engineering Command, Orlando Fl.

[9] Schmidt, D. (2007). Krauss-Maffei Weggman (KMW) Embedded Training Concept for Puma Infantry Fighting Vehicle. Presented to the NATO HFM Exploratory Team 077. September 2007, Munich, Germany.

[10] Schmidt, D. (2009). Embedded Virtual Simulation Research Perspective. Presented to the NATO HFM Research Workshop 169. October 2009, Orlando, Florida.

[11] Sottilare, R. (2009). Making a case for machine perception of trainee affect to aid learning and performance in embedded virtual simulations. NATO Research Workshop (HFM-RWS-169) on Human Dimensions in Embedded Virtual Simulations. Orlando, Florida, October 2009.

[12] Sottilare, R. and Roessingh, J. (2011). The application of intelligent agents in Embedded Virtual Simulations (EVS). In Proceedings of the International Training and Education Conference 2011, Cologne, Germany, May 2011.

[13] Sottilare, R. and Roessingh, J.J. (2014). Chapter 6 – Application of Intelligent Agents in Embedded Virtual Simulations. In NATO Technical Report: Improving Human Effectiveness through Embedded Virtual Simulation: Findings of Task Group HFM-165. *NATO Science & Technology Organization*. ISBN: 978-92-837-0181-1.

[14] Sottilare, R., Holden, H., Brawner, K. and Goldberg, B. (2011). Challenges and Emerging Concepts in the Development of Adaptive, Computer-based Tutoring Systems for Team Training. In Proceedings of the Interservice/Industry Training Simulation & Education Conference, Orlando, Florida, December 2011.

[15] Burke, S., Feitosa, J. and Salas, E. (2015). The unpacking of team models in GIFT. In: Sottilare R, Sinatra A, editors. Proceedings of the 3rd Annual GIFT Users Symposium (GIFTSym3); 2015 June 17-18; Orlando, FL. Aberdeen Proving Ground (MD): Army Research Laboratory (US); c2015. ISBN 978-0-9893923-8-9.

[16] Sottilare, R., Burke, S., Johnston, J., Sinatra, A., Salas, E. and Holden, H. Antecedents of adaptive collaborative learning environments. Presented at the Interservice/Industry Training Simulation and Education Conference; 2015 December; Orlando, FL.

[17] U.S. Army Training & Doctrine Command (TRADOC; 2011). The US Army Learning Concept for 2015 (525-8-2). Washington, DC: Department of the Army.

[18] Air Education and Training Command (AETC; 2008). On learning: The future of Air Force education and training. AETC White Paper, Randolph Air Force Base, Texas.

[19] Baker, D.P., Gustafson, S., Beaubein, J., Salas, E. and Barach, P. (2005). Medical Teamwork and Patient Safety: The Evidence-based Relation. Agency for Healthcare Research and Quality Archived Publication #05-0053. Retrieved May 23, 2016 from http://archive.ahrq.gov/research/findings/final-reports/medteam/index.html.

[20] Kozlowski, S.W.J. and Ilgen, D.R. (2006). Enhancing the effectiveness of work groups and teams. Psychological Science in the Public Interest, 7 (3), 77- 124.

[21] Prichard, J.S., Stratford, R.J. and Bizo, L.A. (2006). Team-skills training enhances collaborative learning. Learning and Instruction, 16, 256-265.

[22] Salas, E., Burke, C.S. and Bowers, C.A. (2001). Team training in the skies: does crew resource management (CRM) training work? *Human Factors*, 43(4), 641-674.

[23] Salas, E., Diaz Granados, D. and Klein, C. (2008). Does Team Training Improve Team Performance? A Meta-Analysis. Human Factors, 50(6), 903-933.

[24] Yammarino, F.J., Mumford, M.D., Vessey, W.B., Friedrich, T.L., Ruark, G.A. and Brunner, J.M. (2014). Collective leadership measurement for the U. S. Army. U. S. Army Research Institute for the Behavioral and Social Sciences Study Report 2014-01. Fort Belvoir, VA. Downloaded July 19, 2016 from http://www.dtic.mil/dtic/tr/fulltext/u2/a596962.pdf.

[25] Caldwell, B.S. (2005). Multi-team dynamics and distributed expertise in mission operations. Aviation, Space, and Environmental Medicine, Jun;76(6 Suppl):B145-53.

[26] Baldwin, E. (2008). Integrating space flight resource management skills into technical lessons for international space flight controller training. United Space Alliance IAASS Conference 2008. Retrieved July 27, 2016 from http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20080041534.pdf.

[27] Marks, M.A., Mathieu, J.E. and Zaccaro, S.J. (2001). A temporally based framework and taxonomy of team processes. Academy of Management Review, 26: 356–376. Retrieved June 8 2016 from http://www.owlnet.rice.edu/~ajv2/courses/12c_psyc438001/Marks%20et%20al.%20(2001).pdf.

[28] Salas, E., Rosen, M.A., Held, J.D. and Weissmuller, J.J. (2009). Performance measurement in simulation-based training: A review and best practices. Simulation and Gaming: An Interdisciplinary Journal, 40(3), 328-376.

[29] Landon, L.B., Vessey, W.B. and Barrett, J.D. (2016). Evidence Report: Risk of Performance and Behavioral Health Decrements Due to Inadequate Cooperation, Coordination, Communication, and Psychosocial Adaptation within a Team. Johnson Space Center, Houston TX. Retrieved July 26, 2016 from https://humanresearchroadmap.nasa.gov/Evidence/reports/Team.pdf.

[30] Carolan, T.F., Bilazarian, P. and Nguyen, L. (2005). Automated individual, team, and multi-team performance assessment to support debriefing distributed simulation-based exercises (DDSBE). Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting.

[31] Stretton, M.L. and Wilson, M.S. (2004). Semi-automated observations and assessment – Trainer interactions within a distributed training environment. Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting.

[32] Radtke, P.H., Johnston, J.H., Biddle, E. and Carolan, T.F. (2007). Integrating and presenting performance information in simulation-based air warfare scenarios. Proceedings of the Interservice/Industry Training, Simulation, and Education Conference.

[33] Pharmer, J.A. and Milham, L.M. (2016). Adaptive training for command-and-control teams: challenges in scenario-based training, Theoretical Issues in Ergonomics Science, 17(2), 169-182, DOI: 10.1080/1463922X.2015.1111461.

[34] Wegner, D.M. (1986). Transactive memory: A contemporary analysis of the group mind. In G. Mullen & G. Goethals (Eds.), *Theories of Group Behavior* (pp. 185-208). New York: Springer-Verlag.

[35] Moreland, R.L. (2006). Transactive memory: Learning who knows what in work. In, J.M. Levine and R.L. Moreland (Eds.), *Small Groups and Organizations: Key Readings* (pp 327-346), New York, NY: Psychology Press.

[36] Lewis, K. and Herndon, B. (2011). Transactive memory systems: Current issues and future research directions. *Organization Science, 22(5)*, 1254-1265.

[37] Briggs, G.E. and Johnston, W.A. (1967). *Team Training* (NAVTRADEVCEN 1327-4). Orlando, FL: Navy Training Device Center. (DTIC/NTIS No. ADA 660 019).

[38] Olmstead, J.A. (1992). *Battle Staff Integration* (IDA Paper P-2560). Alexandria, VA: Institute for Defense Analyses. (DTIC/NTIS No. ADA 248 941).

[39] Salas, E., Shuffler, M.L., Thayer, A.L., Bedwell, W. and Lazzara, E. (2014). Understanding and improving teamwork in organizations: A scientifically based practical guide. Human Resource Management, 54(4). DOI: 10.1002/hrm.21628. Retrieved 6 June 2016 from: https://www.researchgate.net/publication/267571055_Understanding_and_Improving_Teamwork_in_Organizations_A_Scientifically_Based_Practical_Guide.

[40] Landon, L.B., Vessey, W.B. and Barrett, J.D. (2016). Evidence Report: Risk of Performance and Behavioral Health Decrements Due to Inadequate Cooperation, Coordination, Communication, and Psychosocial Adaptation within a Team. *Johnson Space Center*, Houston TX. Retrieved July 26, 2016 from https://humanresearchroadmap.nasa.gov/Evidence/reports/Team.pdf.

[41] 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.

[42] Adamson, D., Dyke, G., Jang, H. and 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.

[43] Rosen, Y. (2015). Computer-based assessment of collaborative problem solving: Exploring the feasibility of human-to-agent approach. *International Journal of Artificial Intelligence in Education*, 25(3), 380-406.

[44] McManus, M.M. and Aiken, R.M. (2016). Supporting Effective Collaboration: Using a Rearview Mirror to Look Forward. International Journal of Artificial Intelligence in Education, 26(1), 365-377.

[45] Sottilare, R.A., Burke, C.S., Salas, E., Sinatra, A.M., Johnston, J.H. and Gilbert, S.B. (2016, *in review*). Towards a Design Process for Adaptive Instruction of Teams: A Meta-Analysis. In R. Sottilare, A. Graesser, J. Lester and R. Baker (Eds.) Special Issue: Generalized Intelligent Framework for Tutoring (GIFT): Creating a stable and flexible platform for innovations in AIED research. *International Journal of Artificial Intelligence in Education.*

[46] Sottilare, R.A., Brawner, K.W., Goldberg, B.S. and 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.

[47] 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.

[48] Woolley, A.W., Bear, J., Chang, J.W. and DeCostanza, A.H. (2011, January). The Effects of Team Strategic Orientation on Team Process in Competitive Environments. In Academy of Management Proceedings (Vol. 2011, No. 1, pp. 1-6). Academy of Management.

[49] Woolley, A.W., Bear, J.B., Chang, J.W. and DeCostanza, A.H. (2013). The effects of team strategic orientation on team process and information search. Organizational Behavior and Human Decision Processes, 122(2), 114-126.

[50] Dillenbourg, P. (1999). What do you mean by collaborative learning? Collaborative-learning: Cognitive and computational approaches, 1, 1-15.

[51] Tegos, S., Demetriadis, S. and Tsiatsos, T. (2014). A configurable conversational agent to trigger students' productive dialogue: A pilot study in the CALL domain. *International Journal of Artificial Intelligence in Education*, 24(1), 62-91.

[52] Drury, H., Kay, J. and Losberg, W. (2003, January). Student satisfaction with groupwork in undergraduate computer science: do things get better? In Proceedings of the *5th Australasian conference on Computing education*. Volume 20 (pp. 77-85). Australian Computer Society, Inc.

[53] Kay, J., Yacef, K. and Reimann, P. (2007, July). Visualisations for team learning: small teams working on long-term projects. In Proceedings of the *8th International Conference on Computer supported collaborative learning* (pp. 354-356). International Society of the Learning Sciences.

[54] 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.

[55] 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.

[56] Sottilare, R., Hackett, M., Pike, W. and LaViola, J. (2016). Adaptive instruction for medical training in the psychomotor domain. *The Journal of Defense Modeling and Simulation*, 1548512916668680.

[57] Goldberg, B. and Amburn, C. (2015). The Application of GIFT in a Psychomotor Domain of Instruction: A Marksmanship Use Case. In Proceedings of the *3rd Annual Generalized Intelligent Framework for Tutoring (GIFT) Users Symposium (GIFTSym3)*.

Chapter 6 – A REVIEW OF INTELLIGENT TUTORING SYSTEMS FOR SCIENCE TECHNOLOGY ENGINEERING AND MATHEMATICS (STEM)

Ray S. Perez, Ph.D., Anna Skinner, Ph.D., and Robert A. Sottilare, Ph.D.

6.1 INTRODUCTION

Robelen [1] indicated that only 16% of American high school seniors are proficient in mathematics and interested in science, technology, engineering, and mathematics (STEM) careers. Even among those who do pursue a college major in the STEM fields, only about half go on to work in a STEM-related career [1]. The United States is falling behind, internationally ranking 25th in mathematics and 17th in science among industrialized nations [1]. Today's approaches to education and training must change if our nation is to be competitive in the dynamic global workforce. Numerous studies have shown that an effective way to teach students is through one-on-one interactions between teachers and students [2]-[4]. One way to achieve one-on-one instruction is through the use of advanced technologies such as Intelligent Tutoring Systems (ITSs).

Lemke [5] argues that learning science subject matter in particular, requires articulation of concepts and conversing with others that have mastered the subject matter and the language for discussing the subject matter.

We describe here recent efforts that represent the state-of-the-art in intelligent tutoring system development, as well as the research gaps and further research and development required to address these limitations in support of science and technology solutions.

6.2 TYPES OF INTELLIGENT TUTORS FOR STEM

The following provides examples of two types of ITS technologies: ACT-R Based Tutors and AutoTutor that are in use in the United States for the most part. These examples are not intended to provide a comprehensive review of existing ITSs (for a detailed review, see Ref. [6]); rather these two tutors are illustrative of how ITS technologies and various instructional strategy implementations have been applied to STEM subject matter. We will describe a tutor for improving literacy skills, reading comprehension and writing. While improving literacy skills is not typically considered as part of STEM learning, it has been pointed out that if you cannot read or write you certainly cannot learn complex STEM concepts [7].

6.2.1 ACT-R-Based Tutors

No review of the literature on ITSs in STEM could be complete without mentioning the Anderson tutors, also known as cognitive tutors [8]. These tutors were designed to teach students to solve problems in physics, geometry, and algebra. Cognitive tutors are in use at least two days per week by more than 600,000 students a year in 2,600 middle or high schools, and represent one of the most widely used ITS technologies. Cognitive tutors have a 20-year history of research and development and are based on the adaptive control of thought (ACT*) theory of learning and problem solving described in *The Architecture of Cognition* [9]. The development of cognitive tutors represents a unique approach to building tutors. They tested the ACT* theory by observing whether the ACT* theory designed tutor could optimize learning. Many of the evaluations of cognitive tutors have found evidence of positive learning effects providing support for the ACT* theory.

However, as previously noted, in some cases these effects have been shown to be relatively small and have been shown to only produce effects after long-term use [10]. The development of ACT-R, a unified computational theory of human architecture, provided a sound scientific foundation for the design of training and education systems.

The underlying architecture of these tutors is driven by the ACT* theory of learning. ACT* theory states that a cognitive skill consists, in large part, of units of goal-related knowledge; and that skill acquisition involves the formulation of thousands of rules that relate to task goals and task state to actions and consequences [8]. This theory also makes a distinction between procedural knowledge and declarative knowledge. Procedural knowledge is the ability to use rules to solve a problem whereas declarative knowledge consists of knowing the concepts and rules. The theory refers to the learning process as knowledge compilations, which convert this interpretative problem solving into production rules, and assumes that production rules can only be learned by employing declarative knowledge within the context of problem-solving activities. The last tenet of this theory is strengthening; much like Thorndike's "law of effect" the tenet suggests that both declarative and procedural knowledge acquired strength with practice, successive practice produces smoother and more rapid execution and fewer errors. Later, Anderson, Boyle, Farrell, and Reiser [11] extracted five principles to guide the design of tutors:

a) Represent student competence as a production set;

b) Communicate the goal structures underlying the problem solving;

c) Provide instruction in the problem-solving context;

d) Promote an abstract understanding of the problem-solving knowledge; and

e) Minimize working memory load.

These tutors make use of two major algorithms: model tracing and knowledge tracing. The model-tracing algorithm evaluates the correctness of each of the student's attempts at solving problems by comparing step by step the student's steps with the possible steps of the cognitive model. The knowledge-tracing algorithm maintains estimates of the probability that the student knows each knowledge component in the cognitive model and is represented by a production rule [12]. The knowledge tracer makes use of the information provided by model tracing to determine when the student has provided a correct action or incorrect action. Anderson and his colleagues initially called their ITSs tutors because they were inspired by the intelligent tutors build in the late 1970s and early 1980s [13]. The initial design goals were intended to enable the tutor to interact with students much like human expert tutors. However, over the 20 or so years of development, the emulation of human tutors has been de-emphasized. The cognitive tutors have been involved in many evaluations of the LISP, geometry, and algebra tutors. The results of the algebra Cognitive Tutor will be presented here as it is widely used [14], and the effects have varied from no significant differences for evaluations that took place in 1987–1988 with the algebra Cognitive Tutor to the most recent wide scale evaluation [10] conducted by the Rand corporation and funded by the U.S. Department of Education, for which small effect sizes were found (i.e., effect size of .22).

The Rand study involved seven states and a variety of middle schools and high schools (including 6,800 students in Grades 6 – 8 and 18,700 students in Grades 9 – 12). One measure of success of ITSs is the prevalence of their use. According to Carnegie Learning Inc., 600,000 students in more than 2,600 middle and high schools use Carnegie Learning's Cognitive Tutor for algebra. Pane *et al.* [10] found the tutor to be more effective than traditional algebra courses. This evaluation matched middle and high school students into pairs randomly assigned to either a traditional algebra curriculum (control) or one that included the algebra Cognitive Tutor (experimental treatment). Positive results in favor of the algebra Cognitive Tutor in the second year of the two-year study were found only for the high schools, but demonstrated that the high school students' performance was superior to the control classes in the second year. Algebra Cognitive Tutor raised their score on average by eight percentile points, for which small effects were found (i.e., effect size of .22).

6.2.2 AutoTutor

AutoTutor is an ITS that helps students learn Algebra 1, Newtonian physics, computer literacy, and critical thinking through the use of advanced tutorial dialogues in natural language. AutoTutor simulates a human tutor by holding a conversation with the learner using natural language processing. The student model of AutoTutor is an example of misconception tailored dialogue [15]. AutoTutor is theory-driven, and much like ACT-R tutors, AutoTutor's design is driven by explanation-based constructivist theories of learning [16]-[18]. In addition to being theory-driven, AutoTutor is an ITS that adaptively responds to the learner's actions. It gives immediate feedback to the tutee and guides the tutee on what do next; this guidance reflects what the tutor infers as the learner's state of knowledge. Tutorial strategies used by AutoTutor are empirically based, grounded in research on strategies that are used by expert human tutors. For example, one of the prominent dialogue patterns used by AutoTutor is Expectation and Misconception Tailored (EMT) dialogue, which has been observed to be used by human tutors [19]-[20]. EMT is based on the observation that human tutors typically have a list of anticipated "good" answers (i.e., expectations) and a list of misconceptions associated with each question or problem, and so does AutoTutor. AutoTutor has at least two goals. The first goal is to coach the student in covering the list of expectations, and the second is to correct the anticipated misconceptions that are detected in the student's interactions with the machine tutor (e.g., talk and actions).

AutoTutor simulates the discourse patterns of human tutors and also incorporates a number of ideal tutoring strategies. Another goal of AutoTutor is to provide feedback (positive, neutral, and negative) that adaptively responds to the student; it pumps the learner for additional information (e.g., "what else?"), prompts the learner to fill in missing words, gives hints, fills in missing information with assertions, identifies and corrects inappropriate answers, answers the learner's questions, and summarizes answers [21]. It does this by presenting a series of challenging problems or questions that require verbal explanations and reasoning in an answer. It also engages in a collaborative mixed-initiative dialogue while constructing an answer; this process typically takes approximately 100 conversational turns. AutoTutor provides the content of its turns through an animated conversational agent using natural language with a speech engine and a sensor that is able to identify some facial expressions and rudimentary gestures. For some topics, there are graphical displays, animations of causal mechanisms, or the use of interactive simulations. AutoTutor tracks the cognitive states of the learner by analyzing the content of the dialogue history. AutoTutor dynamically selects the words and statements in each conversational turn in a fashion that is sensitive to what the learner knows. The AutoTutor system also adapts to learners' emotional states in addition to their cognitive states [21]. The impact of AutoTutor in facilitating the learning of "deep" conceptual knowledge has been validated in over a dozen experiments with college students for topics in conceptual physics [17], [22] and introductory computer literacy [23]. Tests of AutoTutor have produced medium and large effect sizes of 0.4 to 1.5, moves a person from the 50th percentile to the 94th percentile, with a mean of 1.5, depending on the learning measure, the comparison condition, the subject matter, and version of AutoTutor [17].

The latest version of AutoTutor uses shareable knowledge objects that teach Algebra 1 skills with user-friendly authoring tools. Selection of skills will be (vs. was) based on concept maps and learning progression, and Common Core standards [17], [24]. A Shareable Knowledge Objects (SKO) framework declares a composition of services intended to deliver knowledge to a user, with the expected use case being tutoring in natural language. In this context, the SKO framework is not a reimplementation of AutoTutor but a framework for breaking AutoTutor down into minimal components that can be composed to create tutoring modules that may or may not rely on the traditional AutoTutor modules. These minimal components are intended to be used as part of a service-oriented design [25]. This research integrates AutoTutor, a conversational tutoring system, into ALEKS, a commercial K-12 mathematics learning environment used by hundreds of schools. AutoTutor's technology as applied to ALEKS is intended to enhance ALEKS's procedural practice and worked solutions by adding animated and interactive tutoring agents that help students master math concepts. As in the earlier versions [24], tutoring is provided by a pair of animated agents (a computer tutor and a computer student) that talk to the student using natural language; the computer student types to the tutor in free text and the tutor speaks back using voice and also provides text. These tutoring sessions focus primarily on math reasoning, analyzing problems and explaining the underlying concepts. It is expected that these dialogues promote higher learning and motivation within the ALEKS system, as they include animated agents who tutor mathematics concepts. The project has as its goal integrating tutoring into a special ALEKS course covering a subset of the ALEKS Algebra 1 course, which is matched to the U.S.'s initiative of the "Common Core".

More recently, a version of AutoTutor was developed for improving literacy skills using trilogues in which the goal is to help adults with low literacy improve their literacy by engaging students in conversation in natural language interacting with two computer agents that use ideal pedagogical strategies [17]. These tutors using natural language dialogue yield learning gains comparable to those trained by human tutors, with effect sizes ranging from 0.6 to 2.0 [23], [24], [26]. A trilogue is a conversation between a machine tutor and a student in which two agents, one acting as a tutor and the other as a peer, interact with a human student. These agents can take on different roles [17]. These tutors simulate the dialogue between tutee and human tutor with high fidelity, finally approaching the capability for machine tutors to have a conversation with a learner much as envisioned by Jamie Carbonell over 60 years ago; however, these dialogues are scripted.

6.3 ITS INSTRUCTIONAL STRATEGIES FOR STEM

6.3.1 Worked Examples

Providing learners with a sample problem and solution, also referred to as a worked example, following the explicit introduction of one or more domain principles such as a mathematical theorem or physics laws has been shown to enhance learning and transfer of the subject material, and has been shown to be more effective than completion of a sample problem by the learner [27]. Furthermore, empirical studies have demonstrated that learners should be presented with several examples, rather than a single example, in order to increase understanding and generalizability and application of learned concepts to novel problem sets [27].

Much research has been dedicated to examining the most effective instructional strategies involving learning by examples, including incorporation of additional effective learning strategies. For example, Schwonke *et al.* [28] have provided evidence to indicate that the study and self-explanation of worked examples are more likely to have an effect on conceptual learning than on normal learning. Renkl [27] provides a set of guidelines, based on the research literature, for incorporating self-explanation, help, and imagery within worked examples, as well as guidelines for the development of example sets that involve easy mapping, meaningful building blocks, learning by errors, model–observer similarity, and interleaving example study with problem solving by gradually fading worked steps.

6.3.2 Reading and Writing Tutor: Game-Based Learning

Interactive Strategy Training for Active Reading and Thinking (iSTART Writing-Pal) is a game-based ITS that supports the development of literacy skills, specifically for young adult learning and practice of reading comprehension and core writing strategies [29]-[30]. The iSTART reading comprehension tutor was developed based on findings that students who self-explain texts and use effective comprehension strategies are better able to develop deep, coherent understandings of challenging texts. Students in iSTART learn strategies for generating self-explanations through five instructional lessons focusing on comprehension monitoring, paraphrasing, prediction, elaboration, and bridging. During practice, students type self-explanations while reading a text. iSTART provides scaffolded feedback driven by sophisticated computational algorithms. Writing-Pal is a game-based ITS that supports the learning and practice of core writing strategies. Students learn strategies for writing through nine instructional lessons focused on generating and organizing their ideas, drafting persuasive essays with a clear rhetorical structure, and revising their essays to express ideas in a more sophisticated and cohesive manner. Students also practice writing using an essay writing tool, which provides automated formative feedback to guide their overall strategy use and learning.

6.4 RESEARCH GAPS AND FUTURE DIRECTIONS

Today's ITS developers are still faced with the difficult technical tasks of constructing and validating individual student models and developing effective natural language interfaces. Additionally, tutorial dialogues still remain a challenge, as well as diagnosis of students' errors and how to remediate them. Finally, the cost and level of effort required has been daunting. The way forward for improving the performance of intelligent tutoring systems lies in emerging technologies and new pedagogical theories based on the science of learning: virtual human avatars, data mining, leveraging game-based learning, and a value-added approach.

6.4.1 Virtual Human Avatars as an Interface

An important goal of ICAI system developers was to create between the student and the computer a dialogue in much the same way that human expert tutor would interact with a student where questions and answers come from both student and computer [31]. This goal of earlier ITSs achieving the capability to provide a mixed-initiative dialogue was accomplished in a limited way. A promising technology for meeting this goal is development of a sophisticated intelligent virtual human avatar as interface between the tutor and the student. This interface would have not only the capability of speech recognition and generation but also the capability to recognize the meaning of emotional states and gestures. Earlier we described the INOT avatar's limited ability to interact with a trainee using speech recognition and generation to teach interpersonal skills. The avatar in this system is scenario driven and the interactions between the avatar and trainee is completely scripted. For example, if the trainee selects a strategy for resolving a personal problem, the avatar is scripted to follow a pre-specified set of actions based on the trainee's selected strategy. An ITS could be enhanced by building an intelligent avatar that is driven by the domain, learner, and tutor models and is able to provide a mixed-initiative dialogue with trainees that represents real world interactions.

6.4.2 Data Mining to Refine the Student Model

Sleeman and Brown [13] in reviewing earlier ITSs and commenting on their limitations asserted that instructional material produced in response to a student's query or mistake was often at an inappropriate level of detail, as the system assumed too much or too little student knowledge. Recently, there has been a focus on the use of big data for evaluating learning systems and the development and testing of theories that are used to explain learning gains [32]. Koedinger *et al*. [33] have suggested that these data-driven techniques could be used to optimize the selection, evaluation, suggestion, and update functions of intelligent tutors.

Thus, further research on the use of data mining and other data analytics could be used to fine-tune the student model (e.g., learner profiles), ensuring that system generated remediation in response to a student's query or mistake is at the appropriate level of detail and matches the student's level of knowledge. A limitation to this method is that findings are specific to the characteristics of the sample analyzed and only infers from the observed behavioral responses, so we have no idea what students were thinking.

6.4.3 Value-Added Approach to Improve ITS Effectiveness

Kulik and Fletcher [34] in their recent meta-analysis of research on the effectiveness of intelligent tutoring systems found an average effect size of .69 with outliers achieving a 1.5 sigma. The authors also tell us that it is important to have test alignment, that is the tests that are being used to assess the effectiveness of the ITS must be aligned with the learning objectives that the ITS is designed to teach.

The complexity of the topics to be taught is critical; tutors do not do well teaching topics that only require recognition and recall and they are more appropriate for teaching topics that require decision making and problem solving. Finally, the degree of implementation is important. If tutors are to be effective they must be implemented as designed, and teachers need to be taught how to use the tutors appropriately. What this and other evaluations do not indicate is what components of the tutor account for most of the learning. Typical ITS evaluations are at the system level. An evaluation approach is needed that separately evaluates the individual ITS components; this approach has been coined value-added research, proposed by Mayer [35].

This approach is designed to compare the effectiveness of a base system with that of one that has been enhanced (e.g., more robust student model) and measure the added effectiveness (value) of the more robust student model to assess how effective a tutor is if you remove or disable the student model. Only by including those components of a tutor that add value and increase effectiveness might the cost of a tutor be reduced or its effectiveness optimized. The Generalized Intelligent Framework for Tutoring (GIFT) provides a testbed capability to evaluate learner model attributes, instructional strategies, and domain content [36]. Finally, ITSs provide the ability to not only enhance learning, but to provide insight into the learning processes, identifying optimal instructional strategies and providing alternate strategies to the way in which education and training is conducted today.

6.5 RECOMMENDATIONS

This section recommends research and development investments needed for NATO STO to develop or adopt adaptive tutoring capabilities for STEM. Four primary recommendations follow:

- Follow and evaluate best practices and recommendations of the Program for International Student Assessment (PISA), an influential component of the Organization for Economic Cooperation and Development (OECD), to enhance the use of ITSs for collaborative problem solving.
- Conduct a study to identify specific ROIs for STEM training domains that provide near term opportunities) for applying ITS technologies to STEM training.
- Develop ITSs to support STEM topics as needed (e.g., reading, writing, mathematics) to enhance basic STEM skills.
- Evaluate the effectiveness of ITSs developed for STEM as whole systems, at the component level (e.g., learner model), and at the methodology level (e.g., assessment methods).

6.6 REFERENCES

[1] Robelen, E.W. (2012). Public Gets Glimpse of Science Standards. *Education Week*, *31*(31), 1-15.

[2] Bloom, B.S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher, 13*(6), 4-16.

[3] Kulik, J.A. and Fletcher, J.D. (2015). Effectiveness of Intelligent Tutoring Systems A Meta-Analytic Review. *Review of Educational Research*, 0034654315581420.

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

[5] Lemke, J.L. (1990). *Talking science: Language, learning, and values*. Ablex Publishing Corporation, 355 Chestnut Street, Norwood, NJ 07648 (hardback: ISBN-0-89391-565-3; paperback: ISBN-0-89391-566-1).

[6] Woolf, B.P., Lane, H.C., Chaudhri, V.K. and Kolodner, J.L. (2013). AI Grand Challenges for Education. *AI Magazine*, 34(4), 9

[7] McNamara, D.S., Jacovina, M.E., Snow, E.L. and Allen, L.K. (2015). From generating in the lab to tutoring systems in classrooms. *The American journal of psychology*, *128*(2), 159-172.

[8] Anderson, J.R., Corbett, A.T., Koedinger, K.R. and Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences*, 4(2), 167-207.

[9] Anderson, J.R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning* and Verbal Behavior, 22(3), 261-295.

[10] Pane, J.F., Baird, M. and Education, R.A.N.D. (2014). Methods Used to Estimate Achievement Effects in Personalized Learning Schools.

[11] Anderson, J.R., Boyle, C.F., Farrell, R. and Reiser, B.J. (1987). Cognitive principles in the design of computer tutors. In P. Morris (Ed.), *Modeling cognition* (pp. 93-134). New York. Wiley.

[12] Corbett, A.T. and Anderson, J.R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User modeling and user-adapted interaction*, *4*(4), 253-278.

[13] Sleeman, D. and Brown, J.S. (1982). Intelligent tutoring systems.

[14] Koedinger, K.R. and Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review*, *19*(3), 239-264.

[15] Sottilare, R.A., Graesser, A., Hu, X. and Holden, H. (Eds.). (2013). *Design Recommendations for Intelligent Tutoring Systems: Volume 1-Learner Modeling* (Vol. 1). US Army Research Laboratory.

[16] Aleven, V.A. and Koedinger, K.R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science*, *26*(2), 147-179.

[17] Graesser, A.C., D'Mello, S.K. and Strain, A.C. (2014). Emotions in advanced learning technologies. *International handbook of emotions in education*, 473-493.

[18] VanLehn, K., Jones, R.M. and Chi, M.T. (1992). A model of the self-explanation effect. *The journal of the learning sciences*, *2*(1), 1-59.

[19] Graesser, A.C., Hu, X. and McNamara, D.S. (2005). Computerized learning environments that incorporate research in discourse psychology, cognitive science, and computational linguistics. *Experimental Cognitive Psychology and its Applications: Festschrift in Honor of Lyle Bourne, Walter Kintsch, and Thomas Landauer. Washington, DC: American Psychological Association*.

[20] Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied cognitive psychology*, 9(6), 495-522.

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

[22] Graesser, A.C., Wiemer-Hastings, P., Wiemer-Hastings, K., Harter, D., Tutoring Research Group, T.R.G. and Person, N. (2000). Using latent semantic analysis to evaluate the contributions of students in AutoTutor. *Interactive learning environments*, *8*(2), 129-147.

[23] McNamara, D.S., O'Reilly, T.P., Best, R.M. and Ozuru, Y. (2006). Improving adolescent students' reading comprehension with iSTART. *Journal of Educational Computing Research*, *34*(2), 147-171.

[24] Graesser, A.C. and D'Mello, S. (2012). 5 Emotions During the Learning of Difficult Material. *Psychology of Learning and Motivation-Advances in Research and Theory*, *57*, 183.

[25] Nye, B.D., Graesser, A.C. and Hu, X. (2014). AutoTutor and family: A review of 17 years of natural language tutoring. *International Journal of Artificial Intelligence in Education*, *24*(4), 427-469.

[26] Graesser, A.C., Lu, S., Jackson, G.T., Mitchell, H.H., Ventura, M., Olney, A. and Louwerse, M.M. (2004). AutoTutor: A tutor with dialogue in natural language. *Behavior Research Methods, Instruments, & Computers*, *36*(2), 180-192.

[27] Renkl, A. (2011). Instruction based on examples. *Handbook of research on learning and instruction*, 272-295.

[28] Schwonke, R., Wittwer, J., Aleven, V., Salden, R.J.C.M., Krieg, C. and Renkl, A. (2007). Can tutored problem solving benefit from faded worked-out examples. In *Proceedings of EuroCogSci* (Vol. 7, pp. 59-64).

[29] McNamara, D.S., Crossley, S.A., Roscoe, R.D., Allen, L.K. and Dai, J. (2015). A hierarchical classification approach to automated essay scoring. *Assessing Writing*, *23*, 35-59.

[30] Jacovina, M.E., Snow, E.L., Jackson, G.T. and McNamara, D.S. (2015, June). Game features and individual differences: interactive effects on motivation and performance. In *International Conference on Artificial Intelligence in Education* (pp. 642-645). Springer International Publishing.

[31] Carbonell, J.R. (1970). AI in CAI: An artificial-intelligence approach to computer-assisted instruction. *IEEE transactions on man-machine systems*, *11*(4), 190-202.

[32] Scheuer, O. and McLaren, B.M. (2012). Educational data mining. In *Encyclopedia of the Sciences of Learning* (pp. 1075-1079). Springer US.

[33] Koedinger, K.R., Booth, J.L. and Klahr, D. (2013). Instructional complexity and the science to constrain it. *Science*, *342*(6161), 935-937.

[34] Kulik, J.A. and Fletcher, J.D. (2015). Effectiveness of Intelligent Tutoring Systems: A Meta-Analytic Review. *Review of Educational Research*, 0034654315581420.

[35] Mayer, R.E. (2014). Cognitive theory of multimedia learning. *The Cambridge handbook of multimedia learning*, *43*.

[36] Sottilare, R., Goldberg, B., Brawner, K. and 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.

Chapter 7 − A REVIEW OF INTELLIGENT TUTORING SYSTEM TECHNOLOGIES FOR MEDICAL TRAINING AND EDUCATION

Robert A. Sottilare, Ph.D., Ray S. Perez, Ph.D., and Anna Skinner, Ph.D.

7.1 INTRODUCTION

This Chapter examines the application of Intelligent Tutoring System (ITS) technologies (tools and methods) to the medical training domain. ITSs have been used for a variety of medical tasks, but their application in the medical field has room to grow. The most recognizable application of artificial intelligence to the field of medical instruction is in *expert systems*. Expert systems are software programs which use artificial intelligence techniques and databases of expert knowledge to offer advice or make decisions. One of the most prevalent expert systems on the internet is WebMD (see WebMD symptom checker at http://symptoms.webmd.com/default.htm#introView) which allows users to input symptoms and then offers treatment advice and narrows the scope of medical diagnosis. Expert systems differ from ITSs in that ITSs have learner models, instructional models, and assessments of the learner's progress against instructional goals. Expert systems lack these elements and are primarily decision trees focused within a particular domain. In short expert systems are diagnostic tools and ITSs are instructional tools. The marriage of ITSs and expert system technologies show promise for the future of adaptive medical instruction.

7.2 EXAMPLES OF MEDICAL TUTORING TECHNOLOGIES

The next subsections of this chapter review examples of ITSs developed for the medical task domain. The discussion below is divided into three eras or timelines: early (1976 – 1999), current (2000 – 2015), and emerging (2016+) medical tutoring technologies. The list of medical ITSs and related technologies described herein is intended to be illustrative and not exhaustive.

7.2.1 Early Medical Tutoring Technologies

The early medical tutoring technology era includes tools and methods for: medical diagnosis instruction [1], drawing conclusions from diagnostic reasoning [2], how to interpret mammograms [3], and teaching diagnostic reasoning for antibody identification [4].

7.2.1.1 Tools for Medical Diagnosis Instruction

GUIDON, an intelligent computer-aided instructional program for teaching medical diagnosis [1], [5], [6], is arguably the best example of a fully developed Intelligent Tutoring System (ITS) for medical training. Although it has been over thirty years since the inception of GUIDON few applications of intelligent tutoring technologies have been developed to support medical training at the scale of GUIDON. The developers of GUIDON attempted to address many of the critical issues facing developers of ITSs today. GUIDON's origin begins with Shortliffe [7] who had the idea of developing a tutoring program from a knowledge base MYCIN. MYCIN was one of the best-known expert systems at that time and was used to identify bacterial infections. Bacterial infections were often found to be extremely difficult to diagnose so there was also a need to develop instructional methods to expose medical students to the wide range of cases beyond those encountered in their clinical experiences.

Detailed knowledge is needed to build an expert model in complex domains like medicine. In ITSs, expert models are used to assess learner performance by comparing their performance to the performance of an ideal learner or expert. Clancey [6], [8]-[9] had reasoned that could use the MYCIN database to expose medical students to a variety of cases so that they could learn to accurately diagnosis bacterial infections. He began to build a tutorial program from an expert system like MYCIN which contained a large amount of expert knowledge that an author could leverage and thereby avoid the very labor intensive effort of knowledge elicitation from subject matter experts where that knowledge might already exist in expert systems.

An important feature of an expert system is their ability to make recommendations based on input data. They are differentiated from decision support systems which are designed to help clinicians make better decisions rather than actually make recommendations, which is what an expert system does. The recommendations are essentially a prediction of diagnosis or prognosis or prescription (i.e., a treatment recommendation).

MYCIN separated the knowledge base of production rules from the procedural interpreter. This according Wenger [10] enables access to fine-grained, modular pieces of knowledge, which are in the form of declarative statements which can be understood independently. MYCIN had explanation capabilities that could be used to justify the behavior of the system. By providing the capability to trace these chains of inferences, the system reasoning process became explicit and led to its recommended conclusions.

GUIDON was the perfect candidate to extending MYCIN to a knowledge-based tutorial system. Moreover, the motivation for this effort was the characteristics of the expert system. It had a domain–independent infrastructure; and its reasoning engine had been extracted and made into a more generic form which could have then been applied to various task domains. This is still a design goal of ITS developers today. To reduce the cost, time, and skill required to build ITSs, generalized tutoring architectures like the US Army Research Laboratory's Generalized Intelligent Framework for Tutoring (GIFT [11]; see Chapter 3 of this Report) support authoring of tutors primarily through data driven means and reusable components.

There were some important lessons learned from Clancey's efforts that are still relevant today. First, the finding that the design features for an expert system is not the same as for the design of an instructional system. For, example an expert system may have explanation capabilities but in MYCIN they were expressed passively. On the other hand, a tutor must have the capability to provide feedback to actively present knowledge, and to select/present instructional material based on each learner's behavior and needs. Third, ITSs have to be sensitive to each learner's goals in order to guide interaction for an effective learning experience.

GUIDON had some rather ambitious goals of assessing the pedagogical usefulness of MYCIN's knowledge base; understanding the knowledge needed by a tutoring system to be effective; and the development of tutorial strategies in domain-independent terms. Although, the rules in MYCIN were developed for use in an expert system, the developers of GUIDON were able to use the rule-base as a core and enhance it with new capabilities to produce an active tutor. The developers of GUIDON chose the "case method" as a primary pedagogical presentation strategy.

Following from the SCHOLAR teaching program [12] the case method incorporates a mixed-initiative dialog that focuses on the presentation of successive cases to convey MYCIN's knowledge to learners in a realistic problem-solving context. The learner's diagnosis skills are developed by exposing them to cases abstracted from medical archives. This enabled the ITS to focus the dialogue between the tutor and the learner while allowing either the learner or the tutor to initiate conversation. This approach is similar to the Socratic Method used in "How the West Was Won" [13]. GUIDON guides the learner's reasoning through a Socratic dialogue in which the tutor leads a small group in finding the answer to an open question. It uses MYCIN's 450 rules to present concrete examples. In GUIDON, as a case is selected the learner diagnosis is completed by querying the system to obtain important data and allowing the learner to propose a hypotheses.

The learner's behavior is compared to the expert behavior expected and proposed by MYCIN. The tutorial is designed to intervene when the learner asks for help or when their actions are incorrect (e.g., buggy tutoring). For example, GUIDON might suggest to the learner the next steps they should take or advise them that their questions are not relevant to the case being presented. GUIDON had the capability to understand simple sentences and a list of standard commands (e.g., "help", "justify", "summarize"). The feedback to the learner was generated from a store of primitive phrases so in a basic sense GUIDON it was generative; it could produce responses. The learner is allowed to take the initiative to change the topic. This requires the program to respond intelligently by maintaining a record of the context of the interaction (e.g. conditions).

Although GUIDON was viewed as one of the most sophisticated ITSs ever built, it was not as effective as expected. This was largely due to the way its expertise was represented. While the knowledge was accurate and at the right level of granularity, the schema did not reflect the way that expert human diagnosticians actually reasoned so the system's communication incompatible with medical diagnosis practices of the day. Clancey [14]-[16] reconfigured MYCIN to NEOMYCIN, where GUIDON expressed tutoring strategies in terms of an abstract representational scheme. NEOMYCIN separated the reasoning strategy and expressed abstractly tutoring strategies, in terms of tasks that manipulate specific types of information. Clancey built NEOMYCIN with the ability to justify the reasoning processes and interact with users in the terms they could relate to and understand. One of the goals for NEOMYCIN was to develop a generic tutoring paradigm for all of the classes of problem solvers. He defined these classes of problem solvers as heuristic classification. He classified a given case according to a predetermined taxonomy by relating some features of the data to descriptions of the candidate categories. The intention was to make sure that such a system was based on a thorough understanding of the problem–solving process. Clancey also went on to develop HERACLES a generic system for a class of problem-solving strategies that were used to define and create knowledge engineering tools.

In summary, GUIDON was a unique attempt to turn an existing expert system for a complex task into a medical diagnosis tutor. A major contribution was the identification and separate treatment of different types of knowledge that must be made available in order for a tutor to function effectively. Second, the development of GUIDON identified the need to develop an expert system that contained pedagogical expertise. Clancey called this the *tutorial module*. Ideally, the tutorial strategies developed and delivered by the tutorial module should be domain-independent in that they are used to guide learning in a variety of domains. Although, the implementation of the rule-base tutorial module in NEOMYCIN was deemed effective, the use of MYCIN as a source of domain expertise proved to be problematic as its reasoning strategy was inflexible and it had difficulty tracking the learner's progress. One of the lessons-learned from this research revealed some of the limitations of using compiled production systems for instructional purposes. However, the overall result of this project was a new model of diagnostic thinking and it provided the foundation for many of the medical and non-medical (e.g., Science, Technology, Engineering, and Mathematics − STEM) tutors today.

7.2.1.2 Methods for Drawing Conclusions in Diagnostic Reasoning

The method described here is important to ITS development in that it describes an early study where standardized errors were identified and used to judge the performance of the learner as an expert model. The study by Voytovich and Suffredini [2] explored the characteristics of premature diagnostic conclusions within a group of physicians, medical students, and residents. When the study participants were asked to construct complete, precise problem lists from three case abstracts, premature closure occurred frequently, it could be recognized with good inter-rater reliability, and seemed to appear with equal frequency regardless of the level of training. This led to the identification of error patterns in diagnosis reasoning that were classified into four categories: omission, premature closure, incorrect synthesis, and inadequate synthesis. Omission included errors in which the study participant ignored important clinical clues. Premature closure is where the diagnosis of the patient's condition is not justified by the data/facts. Incorrect synthesis is where the conclusions drawn by the physician are contradicted by the data. Finally, inadequate synthesis is where the conclusions that could be supported by data are not drawn.

7.2.1.3 Tools and Methods for Interpreting Mammograms

This paper outlines a conceptual framework for the development of RadTutor, a prototype computer-based tutor which trains radiology residents in the diagnoses of mammograms that indicate breast diseases. The prototype includes:

1) A discussion of the objectives and goals of the radiology residency training program;
2) A review and critique of existing computer-based radiology training environments;
3) A synthesis of an expert-novice study aimed at attaining a cognitive model of problem solving in mammogram interpretation;
4) A description of the results of analyses of authentic radiology resident teaching rounds; and
5) Deriving instructional principles for the design of the mammography tutor [3].

7.2.1.4 Tools for Teaching Diagnostic Reasoning for Antibody Identification

Smith *et al.* [4] reported the results of a study which indicated that, when used by an instructor as a tool to assist with tutoring in a class laboratory setting, use of the Transfusion Medicine Tutor (TMT) resulted in improvements in antibody identification performance of 87 – 93 % ($p <.001$). Based on input from teachers requesting that TMT be designed for use without the presence of an instructor, a new study on the use of TMT without instructor assistance found that performance improved by 64 – 66 % ($p <.001$). These studies demonstrated the use and utility of ITS technologies in both assisted (human instructor in the loop) and unassisted (no human instructor in the loop) modes as might be considered at training centers in NATO countries.

7.2.2 Current Medical Tutoring Technologies

The current medical tutoring technology era includes ITSs for teaching the interpretation of neuroradiological images [17], teaching how to detect diagnostic errors in internal medicine [18], teaching clinical medicine using various media [19], evaluating student modeling techniques [20], training pathologists using natural language [21], and using game platforms to train medical tasks [22]-[23].

7.2.2.1 Teaching the Interpretation of Neuroradiological Images

In 2000, Sharples *et al*. [17] identified a more systematic approach to training, replacing the traditional mixture of ad hoc apprenticeship and formal lectures with a combination of structured tuition and case-based experiential learning. This change in instructional approaches was intended to meet a long standing need to develop general medical knowledge and the development of skills through diagnostic practice. Computer-based practice was developed to allow radiologists to describe images by means of a structured notation for abnormal image features called the Image Description Language, and thereby improve the diagnostic accuracy of general radiologists to equal that of specialists. Toward this end, a Magnetic Resonance (MR) Tutor was developed and evaluated [17]. The MR Tutor was shown to provide a self-study approach to acquiring knowledge and skill in radiological diagnosis. The method of indexing images by lesion position, appearance, and diagnosis along with the overview plot and the structured approach to training and reporting were well received, but their effectiveness relative to previous methods remains an open question.

7.2.2.2 Teaching the Detection of Diagnostic Errors in Internal Medicine

Graber, Franklin and Gordon [18] reviewed 100 cases of diagnostic error involving internists. The errors were identified through autopsy discrepancies, quality assurance activities, and voluntary reports. Each case was evaluated to identify both the system-related and cognitive factors underlying these errors. The goal of the study was to determine the relative influence of system-related and cognitive factors to diagnostic error and to develop a comprehensive working taxonomy. This study found 90 cases involving injury with 33 cases resulting in death. The underlying factors were classified into three categories: no fault, system-related, and cognitive. Seven cases contained only no-fault errors. In the remaining 93 cases, they identified 548 different system-related or cognitive factors (5.9 per case). System-related factors contributed to diagnostic error in 65% of the cases and cognitive factors in 74%. Common system-related factors involved problems with policies and procedures, inefficient processes, teamwork, and communication. Common cognitive problems involved faulty synthesis. The failure to continue considering reasonable alternatives after an initial diagnosis was the single most common problem. Other common problems included faulty context generation, misjudging the salience of findings, faulty perception, and errors arising from the use of heuristics. Faulty or inadequate knowledge was uncommon. The conclusions of this study commonly involved more than one factor or problem and usually involved both system-related and cognitive factors. This information was be used to classify common diagnostic errors which could be detected by a tutoring system. Today, we regularly identify and model learner errors and misconceptions as part of domain ontology.

7.2.2.3 Teaching Clinical Medicine Using Various Media

Medical students of the University of Ulm use a web-based, case-oriented tutoring system called “Docs ‘n Drugs − The Virtual Polyclinic” [19]. This ITS includes three models:

- A tutoring process model;
- A case knowledge model; and
- Medical knowledge model.

The tutoring process model is roughly equivalent to an instructional model in most tutoring systems. The case knowledge and medical knowledge models together are equivalent to a domain model in most tutoring systems. The tutoring process is described as a series of nodes and steps that describe the structure of medical cases and knowledge which form the expert model within the tutor’s domain model. This approach allows learners to acquire and use knowledge to construct mental models of the medical domain and then apply their skills in case-based environments. This approach is becoming more common in domains where knowledge is well-defined, but its application is less well-defined and grows with experience.

7.2.2.4 Evaluating Student Modeling Techniques

Yudelson, Medvedeva, and Crowley [20] developed a multifactor approach to evaluating student models for a medical ITS which might be broadly applied to other ITSs in other task domains. This approach improved upon the common Bayesian Knowledge Tracing by:

1) Devising a method for evaluating and selecting student models based on decision theoretic evaluation metrics;
2) Evaluating the model selection methodology against independent student performance data; and
3) Investigating the effect of key variables on model performance in a medical diagnostic task.

The combination of areas under the receiver-operator curve (specificity vs. sensitivity) and the precision-recall curve were found to be a valid method for model selection.

7.2.2.5 Training Pathologists Using Natural Language

El Saadawi *et al.* [21] developed and evaluated a Natural Language Interface (NLI) for an ITS used to teach diagnostic pathology to medical residents. The goal of the ITS is to instruct residents on proper methods to examine pathologic slides and write accurate pathology reports. The ITS provided real-time feedback on the resident's errors during their slide review and diagnostic reports. The residents were able to ask for help at any point in the review and received context-specific feedback through natural language. This affiliated study conducted using this ITS revealed good performance with respect to other NLIs at the time in spite of the highly complex and precise language required to describe each pathology slide. The instruction resulted in a highly significant improvement in report writing after only one tutoring session. This study illustrated the effective use of NLIs in ITSs for highly specialized task domains and demonstrates the ability to apply NLIs to highly specialized training domains in NATO countries.

7.2.2.6 Using Existing Game Platforms to Train Medical Tasks

In 2012, the US Army Research Laboratory undertook the challenge to use existing game platforms to tutor the cognitive aspects of medical tasks. Specifically, the Generalized Intelligent Framework for Tutoring (GIFT) was linked to both the Virtual Battlespace 2 (VBS2) and Virtual Medic (VMedic) serious games to drive adaptive instruction for casualty care under fire and hemorrhage control tasks respectively. This demonstrated the ability to assess progress toward task goals based on information in an external environment and enabled GIFT to push adaptive feedback into each serious game in both voice and text modalities.

7.2.3 Emerging Medical Tutoring Technologies

The findings reported in this section were primarily derived from NATO RTG HFM-237's technical review of the medical task domain. Most of these findings include technologies which have been published very recently or are pending publication. The emerging medical tutoring technology era includes ITSs for: tutoring medical psychomotor tasks in the wild [22], modeling skill decay for surgical tasks [23], and developing an ITS for robot-assisted laparoscopic surgery [24].

7.2.3.1 Tutoring Medical Psychomotor Tasks in the Wild

A primary goal of this research is to develop methods to allow tutoring to take place beyond classroom and desktop applications. Toward this end, Sottilare, Hackett, Pike, and LaViola put forth a concept for the adaptive instruction of hemorrhage control tasks in areas without major training infrastructure (in the wild). Their concept involved the use of smart glasses to allow visualization layers on a medical mannequin and hands free training. Along with pressure sensors that were embedded in tourniquets and pressure bandages, this concept provides mechanisms to monitor pressure on virtual wounds, determine blood loss rates based on tourniquet or bandage pressure, and visualize blood flow from wounds. While this concept has not yet been built, it was documented [22] and determined to be a feasible GIFT tutor. A prototype of this concept will be built soon.

7.2.3.2 Modeling Skill Decay for Surgical Tasks

Skinner [23] presented a task analysis for military medicine skill decay. Specifically, the task analysis focused on skill decay related to laparoscopic skill decay from a psychomotor, cognitive, and perceptual point of view. Skinner noted that skill decay presents a challenge across domains during periods of non-use follow initial training and that periods of non-use are common in the military. Specialized skills such as Laparoscopic Surgery are susceptible to decay during deployments. Training gaps were also noted and these included a lack of sensitive and objective metrics, a lack of integrated skill assessments, and a lack of retraining strategies.

An investigation of approaches to overcome these gaps was reported in the results of two experiments. The first experiment focused on methods to acquire fundamental skills in laparoscopic surgery and collected data on both short term (< 6 weeks) and long term retention (approximately 7 months). Laparoscopic surgery requires the use of both hands with equal dexterity. The second experiment built on the first and assessed relative skill decay for psychomotor, cognitive, and integrated tasks, and the comparative effectiveness of video-based refresher training for integrated task skills. The full results of these experiments are pending, but initial results are promising. For the three week retention assessment: the greatest time-based decay was for cognitive tasks; the greatest accuracy-based decay was for integrated tasks; and there was no significant decay for psychomotor task.

7.2.3.3 An ITS for Robot-Assisted Laparoscopic Surgery (RALS)

Kumar [24] reported on the development of an ITS to support tutoring of robot-assisted laparoscopic surgery. The goal was to provide individualized instruction comparable to an expert human instructor by integrating ITS technologies with a simulated RALS surgical procedure. Kumar is planning to conduct usability and training transfer studies to evaluate the efficacy of the ITS. The usability test will study and refine the user experience, and the training transfer study will evaluate near and far transfer of tasks from the practice environment to the operational environment.

7.3 REFERENCES

[1] Clancey, W.J. (1983). Knowledge-based tutoring: The GUIDON Program. *Journal of Computer-Based Instruction*. 10.8-14.; 1986.

[2] Voytovich, A.E., Rippey, R.M. and Suffredini, A. (1985). Premature conclusions in diagnostic reasoning. *Academic Medicine*, 60(4), 302-7.

[3] Azevedo, R. and Lajoie, S.P. (1998). The cognitive basis for the design of a mammography interpretation tutor. In *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society*: August 1-4, 1998, University of Wisconsin-Madison (Vol. 20, p. 78). Psychology Press.

[4] Smith, P.J., Obradovich, J.H., Guerlain, S.A., Rudmann, S., Strohm, P., Smith, J.W. and Sachs, L. (1998, August). Successful use of an expert system to teach diagnostic reasoning for antibody identification. In *International Conference on Intelligent Tutoring Systems* (pp. 354-363). Springer Berlin Heidelberg.

[5] Clancey, W.J. (1986). The science and engineering of qualitative models. Technical report, Stanford knowledge systems laboratory, Stanford University.

[6] Clancey, W.J. (1987*)*. Knowledge-based tutoring: the GUIDON program. *MIT press.*

[7] Shortliffe, E.H. (1976). Computer-based Medical Consultations: MYCIN. New York: *Elsevier/North-Holland Inc.*

[8] Clancey, W.J. (1979). Tutoring rules for guiding a case method dialogue. *International Journal of Man-Machine Studies*, *11*(1), 25-49.

[9] Clancey, W.J. (1983). Communication, Simulation and Intelligent Agents: Implications of Personal Intelligent Machines for Medical Education. *Department of Computer Science, Stanford University*.

[10] Wenger, E. (2014). *Artificial intelligence and tutoring systems: computational and cognitive approaches to the communication of knowledge*. Morgan Kaufmann.

[11] Sottilare, R.A., Brawner, K.W., Goldberg, B.S. and 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.

[12] Carbonell, J.R. (1970). Mixed-Initiative Man-Computer Instructional Dialogues. Final Report 1971. *Bolt, Beranek, and Newman, Boston, Massachusetts.*

[13] Burton, R.R., and Brown, J.S. (1976). A tutoring and student modelling paradigm for gaming environments. *ACM SIGCUE Outlook*, 10 (SI), 236-246.

[14] Clancey, W.J. (1984). Methodology for building an intelligent tutoring system. *Methods and tactics in cognitive science*, 51-84.

[15] Clancey, W.J. (1984). *Classification problem solving* (pp. 49-55). Stanford, CA: Stanford University.

[16] Clancey, W.J. (1986). From GUIDON to NEOMYCIN and HERACLES in twenty short lessons. *AI Magazine*, *7*(3), 40.

[17] Sharples, M., Jeffery, N.P., du Boulay, B., Teather, B.A., Teather, D. and du Boulay, G.H. (2000). Structured computer-based training in the interpretation of neuroradiological images. *International journal of medical informatics*, *60*(3), 263-280.

[18] Graber, M.L., Franklin, N. and Gordon, R. (2005). Diagnostic error in internal medicine. *Archives of internal medicine*, *165*(13), 1493-1499.

[19] Martens, A., Bernauer, J., Illmann, T. and Seitz, A. (2001). "Docs 'n drugs--the virtual polyclinic": an intelligent tutoring system for web-based and case-oriented training in medicine. In *Proceedings of the AMIA Symposium* (p. 433). American Medical Informatics Association.

[20] Yudelson, M.V., Medvedeva, O.P. and Crowley, R.S. (2008). A multifactor approach to student model evaluation. *User Modeling and User-Adapted Interaction*, *18*(4), 349-382.

[21] El Saadawi, G.M., Tseytlin, E., Legowski, E., Jukic, D., Castine, M., Fine, J., and Crowley, R.S. (2008). A natural language intelligent tutoring system for training pathologists: implementation and evaluation. *Advances in health sciences education*, *13*(5), 709-722.

[22] 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).*

[23] Skinner, A. (2015). Surgical Skill Decay and Multimodal Task Analysis. Presented to the Human Factors and Medicine Panel's Research Task Group 237. Santa Margherita, Italy. September 2015.

[24] Kumar, R. (2015). An Intelligent Tutoring System for Robot-assisted Laparoscopic Surgery. Presented to the Human Factors & Medicine Panel's Research Task Group 237. Santa Margherita, Italy. September 2015.

Chapter 8 − SUMMARY AND RECOMMENDATIONS FOR THE EXPLOITATION OF ITS TECHNOLOGIES IN NATO

Robert A. Sottilare, Ph.D.

8.1 INTRODUCTION

This Chapter summarizes Intelligent Tutoring System (ITS) capabilities, gaps, and challenges, and provides a set of recommendations discussed in detail in the proceeding chapters of this report. Finally, a global listing of ITS research and development activities and scientists is provided for reference. The goal of this report is to expose decision makers in NATO countries to the capabilities, potential, and limitations of ITS technologies (tools and methods), and to make them aware of research areas needed to optimize the potential of ITSs as learning tools.

8.2 ITS CAPABILITIES, GAPS, AND CHALLENGES

This Report introduced HFM-237's work plan for investigating ITS capabilities and opportunities to exploit ITS technologies for NATO use. We provided a historical perspective of ITS development and potential return on investment, ROI (Chapter 2), and discussed emerging ITS capabilities (Chapter 4). The Research Task Group (RTG) also reviewed existing ITS capabilities applied to the authoring of ITSs (Chapter 3), collective (team) tutoring domains (Chapter 5), science, technology, engineering, and mathematics (STEM) education (Chapter 6), and medical training and education (Chapter 7). The references within these chapters also provide the reader with a handy list of past and present research in the task domains and topics reviewed by the RTG.

The application of ITS technologies to educational and training task domains varies in maturity, breadth, and depth. ITS technologies have been primarily applied to well-defined domains like mathematics, physics, and software programming where there is usually one (or few) paths to success. It is in well-defined domains for individual learners that ITSs have the highest level of maturity. While mathematics, physics, and software programming may not be tasks frequently undertaken in the military, there are lessons to be learned from the design and development of these tutors that may be transferred to other domains for individual learning. Procedural military tasks (e.g., land navigation, marksmanship, and combat casualty care) are proving to be "tutorable" task domains. As noted earlier, ITSs have been applied primarily to cognitive tasks (e.g., problem solving and decision making) for individuals in well-defined domains. However, there are two areas which NATO should focus resources and effort to optimize the utility of ITSs for military training and education: ITSs in psychomotor domains and ITSs for teams.

First, while there continues to be gaps in ITS capabilities for individual learners (e.g., highly accurate learner state assessments), the major challenges to the practical use of ITSs in the military include the development of tools and methods to train psychomotor tasks involving physical coordination and skill, and conducted beyond desktop computer applications or classroom environments (also known as training in the wild). NATO should endeavor to discover methods to allow tutoring of military members in environments more closely aligned with the operational environment, and thereby support higher levels of skill transfer from training to operations. The primary gap in realizing this capability is the sensing and classification of individual learner behaviors.

Second, since most military operations involve teams and teamwork, NATO should endeavor to discover methods to author, deliver, and automatically manage adaptive instruction of teams. The primary gaps associated with this capability are:

1) Concurrent modeling of individual behaviors and progress toward individual and team objectives;
2) Optimal selection of instructional strategies; and
3) Assessment of teamwork skills (e.g., coaching and conflict management).

8.3 SUMMARY OF RECOMMENDATIONS

Recommendations have been provided in the application chapters (5, 6, and 7) for those decision makers specifically interested in collective (team) training and education, STEM education, and medical training and education. This section summarizes the RTG's recommendations regarding the exploitation of existing ITS capabilities and investments in research. Major recommendations fall into four primary categories, as detailed in the subsections below.

8.3.1 Authoring Tool Recommendations

Authoring tools for specific domains have improved substantially (especially in cognitive domains), but investment is needed to:

- **Expand ITS Authoring Tool Domains** – authoring tools should be expanded to support development of ITSs beyond cognitive tasks (e.g., problem solving and decision-making). Expanding ITS authoring tools to support development of ITSs for affective tasks (e.g., moral dilemmas), psychomotor tasks (e.g., marksmanship), and collective tasks (e.g., team tasks and collaborative problem solving) will enhance the relevance of ITSs to military training and educational domains.
- **Enhance User Authoring Experiences (UX)** – develop, evaluate, and integrate methods to enhance user authoring experiences to reduce the workload, skill and time needed to author effective ITSs.
- **Enhance Authoring Automation** – develop methods to automate authoring tasks (e.g., scenario development and content organization; ITS design, development, deployment, and maintenance; and integration with sensors and external systems (serious games and simulations) to the maximum extent possible and thereby reduce authoring workload and improve the return-on-investment for ITSs to a level where they become a ubiquitous solution for NATO training and education.

Authoring tools is the single most important area of challenge and impact. HFM-237 highly recommends investments be made by NATO countries in ITS authoring tools to have the most impact on ITS affordability and usability. Without a focus on ITS authoring tools, ITSs will continue to be difficult and costly to develop, and impractical for widespread NATO use.

8.3.2 Standardization Recommendations

Another important area of challenge and impact is the pursuit of ITS standards. The development of standards to promote ITS reuse and interoperability will have a major impact on the affordability of ITSs. While nearly all ITSs have models for learners, instruction (pedagogy), and domains, these models and their interaction differ from tutoring system to tutoring system, and do not provide opportunities for reuse. Development of standards and interoperability among ITSs would promote the reuse of ITS frameworks, models, and components. Increasing interoperability will ease the integration of ITSs with other ITSs, existing training infrastructure and educational systems (e.g., edX), and reduce authoring time and cost. Currently, there are no standard processes, models or components for ITSs. Candidates for ITS standardization include:

- ***Learner Modeling*** − We recommend the development of standard learner model attributes which include both domain-independent (e.g., demographics) and domain-dependent (e.g., domain competency, past performance and achievements) fields which are populated from a learner record store (LRS) or long-term learner model. This will promote standard methods to populate real-time models during ITS-based learning experiences and allow for common open learner modeling approaches and transfer of competency models from one tutor to another.

- ***Instructional Strategies*** − We recommend development of standard instructional policies (e.g., mastery learning, faded worked examples, error-sensitive feedback) based on best instructional practices from empirical data on their effect found in the literature. The pedagogical module in the US Army's Generalized Intelligent Framework for Tutoring (GIFT) is based on domain-independent instructional strategies which may be used across training task domains. We recommend leveraging GIFT, an open ITS architecture, and its components to the greatest extent possible.
- ***Inter-Module Message Sets*** − We recommend the development of standard messages to support a common message set for communicating between the four most prevalent components in ITSs (learner, instructional, domain, and tutor-user interface). This will allow various ITS frameworks to interoperate or to repurpose components of one ITS in others to optimize effectiveness. GIFT has standard message sets to facilitate communication between its modules. We recommend leveraging GIFT, an open ITS architecture, to the greatest extent possible.
- ***Experience Application Programming Interface (xAPI)*** − The xAPI is an e-learning software specification that allows learning content and learning systems to speak to each other in a manner that records and tracks all types of learning experiences (e.g. education, training, job experience, reading). Learning experiences are recorded in an LRS or long term learner model. We are recommending standardization of ITSs to be able to produce and consume xAPI statements as a methodology for understanding learner and team domain competency.

Additional ITS standardization recommendations might grow from a formal NATO action group tasked to address development of a standardization agreement for ITSs.

8.3.3 Data Analytic Recommendations

Methods are needed to capture and use data from adaptive instructional experiences in ITSs to:

- ***Model Individual Learners, Teams and Populations*** – Research is needed to support optimal adaptive instructional decisions by ITSs by enhancing the accuracy of models of various states and traits of learners, teams, and populations.
- ***Measure Effectiveness*** – We recommend research and development of a data mining system to understand the effectiveness of adaptive instructional decisions and ITS components to optimize ITS courses and programs of instruction.

8.3.4 Adaptive Interface Recommendations

Tailored information is needed by ITS users depending on their roles in the ITS development, delivery, and maintenance process. We recommend user dashboards be developed to support various users in the ITS user chain: learners, instructors/teachers, instructional designers, ITS authors, ITS researchers, and training/course managers.

8.4 GLOBAL ITS RESEARCH AND DEVELOPMENT ACTIVITIES

This section provides a review of various ITSs scientists around the world and the focus of their research. This list of ITS scientists and associated laboratory programs were solicited via the Society for Artificial Intelligence in Education (AIED) or found through open sources on the internet (refer to Annex A). The inclusion of organizations and affiliated scientists in this report does not infer any official endorsement by the author or their agency. The author's goal was to provide a list as fully representative of those working in the field of ITS research and/or adaptive instructional technologies as possible. The listing in Annex A is not intended to be exhaustive, but is intended to support future cooperation in the development of ITS capabilities.

Below is a listing of countries and the number of scientists engaged in ITS development and/or the discovery of adaptive instructional tools and methods, along with active links hyperlink to available biographies in Annex A:

- Australia (AUS) – 1: Judy Kay.
- Austria (AUT) – 1: Dietrich Albert.
- Brazil (BRA) – 1: F.N. Akhras; Seiji Isotani.
- Canada (CAN) – 9: Cristina Conati; Michel Desmarais; Jim Greer; Ming Hou; Vivekanandan Kumar; Suzanne Lajoie; André Mayers; Ido Roll; Julita Vassileva.
- China (CHN) – 2: Xiangen Hu; Ronghuai Huang.
- Croatia (HRV) – 3: Ani Grubišić; Slavomir Stankov; Branko Žitko.
- Czech Republic (CZE) – 1: Radek Pelanek.
- France (FRA) – 3: Monique Grandbastien; Nathalie Guin; Stéphanie Jean-Daubias.
- Germany (DEU) – 2: Thomas Alexander; Wolfgang Schnotz.
- Greece (GRC) – 1: Ioannis Hatzilygeroudis.
- India (IND) – 1: Shailaja Sardessai.
- Israel (ISR) – 1: Rachel Or-Bach.
- Japan (JPN) – 2: Riichiro Mizoguchi; Kazuhsia Seta.
- Korea (KOR) – 1: Key Sun Choi.
- Luxembourg (LUX) – 1: Samuel Greiff.
- Malaysia (MYS) – 2: Alicia YC Tang; Choo-Yee Ting.
- Mexico (MEX) – 2: Leonardo Garrido; Ramon Zatarain-Cabada.
- Netherlands (NLD) – 1: Bert Bredeweg.
- New Zealand (NZL) – 1: Tanja Mitrovic.
- Philippines (PHL) – 2: Mercedes Rodrigo; Merlin Suarez.
- Portugal (PRT) – 1: Ana Paiva.
- Spain (ESP) – 3: Jesus Gonzalez Boticario; Mikel Larranaga; Olga Santos.
- Sweden (SWE) – 1: Jo-Anne Baird.
- Switzerland (CHE) – 1: Pierre Dillenbourg.

- Thailand (THA) – 1: Nilubon Tongchai.
- Taiwan (TWN) – 2: Tak-Wai Chan; Bor-Chen Kuo.
- United Kingdom (GBR) – 7: Jo-Anne Baird; Benedict du Boulay; Susan Bull; Dragan Gasevic; Johanna Moore; Richard Noss; John Self.
- United States of America (USA) − 97: Vincent Aleven; Ivon Arroyo; Ryan Baker; Tiffany Barnes; Chris Berka; Gautam Biswas; Stephen Blessing; Amy Bolton; Keith Brawner; Peter Brusilovsky; C. Shawn Burke; Winslow Burleson; William Clancey; Joseph Cohn; Ron Cole; Nancy Cooke; Mark Core; Brandt Dargue; Chris Dede; Sidney D'Mello; Michael Dorneich; Paula Durlach; Stephen Fiore; Dexter Fletcher; Peter Foltz; Kenneth Forbus; Jared Freeman; Libby Gerard; Stephen Gilbert; Benjamin Goldberg; Ilya Goldin; Robert Goldstone; Gregory Goodwin; Jamie Gorman; Arthur Graesser; Glenn Gunzelmann; Jiangang Hao; Neil Heffernan; Xiangen Hu; Tanner Jackson; Lewis Johnson; Joan Johnston; Irvin Katz; Kenneth Koedinger; Rohit Kumar; Michelle LaMar; David Landy; Chad Lane; Walter Lasecki; Joseph LaViola; Douglas Lenat; Alan Lesgold; James Lester; Marcia Linn; Noboru Matsuda; Camillia Matuk; Danielle McNamara; Piotr Mitros; Jason Moss; Bradford Mott; Kasia Muldner; Tom Murray; Rodney Nielsen; Benjamin Nye; Brent Olde; Andrew Olney; Philip Pavlik; Sandy Pentland; Ray Perez; Octav Popescu; Anna Rafferty; Sowmya Ramachandran; Mark Riedl; Steven Ritter; Robby Robson; Carolyn Rose; Vasile Rus; Eduardo Salas; Sae Schatz; Dylan Schmorrow; Christian Schunn; David Shaffer; Valerie Shute; Anne Sinatra; Robert Slavin; Robert Sottilare; Gerry Stahl; Ronald Stevens; David Traum; Kurt VanLehn; William Vessey; Wayne Ward; Joseph Williams; Eliot Winer; Beverly Woolf; Michael Young; Diego Zapata-Rivera.

8.5 ACKNOWLEDGMENTS

HFM-237 wishes to recognize the hard work and contributions of its members in the development of this final report. The efforts and expertise of our members provided a basis for our success. Our membership included: Dr. Robert Sottilare (USA − Army Research Laboratory), Dr. Ray Perez (USA − Office of Naval Research), Dr. Thomas Alexander (DEU – FKIE), Dr. Vladimir Zotov (CAN − Defence R&D Canada), Dr. Scott Douglass (USA − Air Force Research Laboratory), Dr. Paula Durlach (USA − Army Research Laboratory), Dr. Dexter Fletcher (USA − Institute for Defense Analyses), Professor Paul Noja (ITA – University of Genoa), Ms. Susan Sidey and Mr. Andrew Green (GBR – Defence Science & Technology Laboratory).

A special thank you goes to our group advisor, Mr. Paul Chatelier (USA − Naval Postgraduate School), and our consultant, Dr. Anna Skinner (USA).

We also wish to acknowledge the hosts and many guest speakers who contributed to the knowledge base and recommendations in this report: LtCol Ron Verkerk (NATO), Dr. Julie Stark (USA – ONR Global London), Professor Arthur Graesser (USA − University of Memphis), Dr. Branko Žitko and Dr. Ani Grubišić (HRV – University of Split), Dr. Bruce Roberts and Dr. Rohit Kumar (USA - Raytheon BBN), Dr. Shawn Burke (USA − University of Central Florida), Dr. Benjamin Goldberg, Dr. Heather Holden, and Dr. Joan Johnston (USA − Army Research Laboratory), Dr. Stephen Gilbert (USA – Iowa State University), Professor Bert Bredeweg (NLD – University of Amsterdam), Dr. Sven Fuchs (DEU − FKIE), Dr. Johanna Moore and Dr. Myroslava Dzikovska (GBR – University of Edinburgh), Dr. Neil Smith (GBR – Defence Science and Technology Laboratory), Dr. Jan Joris Roessingh, Dr. Armon Toubman, and Ms. Pia Justen (NLD – Netherlands Aerospace Centre – NLR), Mr. Joe Armstrong, Dr. Stuart Grant, Dr. Ming Hou, and Dr. Aidan Thompson (CAN – DRDC).

Annex A – INTELLIGENT TUTORING SYSTEM SCIENTISTS

Robert A. Sottilare

This annex provides a list of scientists who are researching and developing ITSs, adaptive instruction and ITS-related technologies (tools and methods) which will influence the future of adaptive instruction. In particular, these scientists are defining architectures, ontologies, and processes for: authoring ITSs; modeling individual learners and teams of learners; automatically managing delivery and presentation of adaptive instruction; modeling domains including expert models and learning assessment methods; and developing effectiveness evaluation methods and testbeds. This listing is presented in alphabetical order based on the scientist's last name and their place of work is indicated by country.

Prof. Dr. Deitrich Albert (AUT) is currently the Head of Cognitive Science Section and Senior Scientist at Knowledge Technologies Institute (KTI) at the Technical University of Graz. Prof. Albert graduated from the University of Göttingen (Germany) with a degree in psychology. He received his Dr.rer.nat. and his postdoctoral degree (Habilitation en Psychologie) from the University of Marburg/Lahn (Germany). He was Professor of experimental psychology (Allgemeine Experimentelle Psychologie) at the University of Heidelberg, Germany. Currently he is senior scientist at TU Graz (KTI), professor emeritus of psychology at University of Graz and key researcher at the Know-Center. His research topics cover several areas in experimental and applied psychology; his actual research focuses on knowledge and competence structures, decision-structures, their applications and empirical research. He is (co-) editor of e.g. three books on knowledge structures (Springer Verlag & Lawrence Erlbaum Ass.). He was the Chair of the Board of Trustees of the Center for Psychological Information and Documentation (ZPID), Germany; and he is a member of several scientific advisory boards. His expertise in European and national R&D projects is documented several grants. The focus of research within these projects has been on modeling knowledge and competence, on developing comprehensive knowledge representation frameworks and ontologies, and on their application in Technology Enhanced Learning (TEL) for personalizing learning paths, adaptively assessing knowledge and competencies of learners, personalized teaching, self-regulated learning, game-based learning and personalized learning environments. Research in decision-making focuses on cognitive processes of analysts and their cognitive biases, and decision support in disaster management. Furthermore, he aims on the evaluation of adaptive systems from a psychological point of view.

Dr. Vincent Aleven (USA) is an Associate Professor in Carnegie Mellon University's (CMU) Human-Computer Interaction Institute, and has over 20 years of experience in research and development of advanced learning technologies, such as ITSs and educational games. Major themes in his research are self-regulated learning, metacognition authoring tools, and the use of tutoring technology in ill-defined domains. Dr. Aleven and colleagues created the Cognitive Tutor Authoring Tools (CTAT), a suite of efficient, easy-to-learn, and easy-to-use authoring tools for intelligent tutoring systems (http://ctat.pact.cs.cmu.edu), including a new paradigm called "example-tracing tutors" that make tutor authoring 4–8 times as cost-effective. CTAT tutors have been built for a wide range of domains, including mathematics (at the elementary school, middle school, and high school level), science (chemistry, genetics), engineering (thermodynamics), language learning (Chinese, French, and English as a Second Language), and learning of intercultural competence. Dr. Aleven is a member of the Executive Committee of the Pittsburgh Science of Learning Center (PSLC), a National Science Foundation (NSF)-sponsored research center spanning CMU and the University of Pittsburgh. He is a co-founder of Carnegie Learning, Inc., a Pittsburgh-based company that markets Cognitive Tutor™ math courses. He was the program committee co-chair of the 2010 International Conference on Intelligent Tutoring Systems. He is co-editor in chief of the *International Journal of Artificial Intelligence in Education*. He has been or is principal investigator (PI) on 7 major research grants and co-PI on 10 others. He has over 200 publications to his name.

Dr. Thomas Alexander (DEU) is head of the Research Group Human Factors of the Ergonomic and Human Machine Systems department of Fraunhofer FKIE. He has received his Ph.D. degree in Engineering at the University of Wuppertal, Germany, and has been working in the field of human factors and ergonomics since 1994. During his time at FKIE, Dr. Alexander has conducted and led many research and technology projects in the related areas. Dr. Alexander has published more than 100 papers and chairs several national and international technical committees and boards in the domain of modeling and simulation of human factors. At present, Dr. Alexander is chairman of the Technical Committee on Human Simulation and Virtual Environments of the International Ergonomic Association (IEA). He also represents German interests within different research groups of the NATO Research and Technology Organisation (RTO) and the European Defense Agency. Together with NATO-SCI 176 Dr. Alexander has received the NATO RTO Scientific Achievement Award in 2010. His main research interests and activities are human behavior modeling, virtual and augmented environments with a special focus on human factors and training. Dr. Alexander is currently a member of NATO Human Factors and Medicine Panel's HFM-237 Research Task Group on the "Assessment of Intelligent Tutoring System Technologies and Opportunities".

Dr. Ivon Arroyo (USA) is Assistant Professor in Learning Sciences and Technologies at Worcester Polytechnic Institute, Worcester, MA. She has designed and implemented tutoring systems for mathematics education, and carried out research on how students learn and perceive mathematics with intelligent tutors at the K-12 level in public school settings. She is PI or Co-PI of NSF and U.S. DoEd grants that predicted student emotions in real time and designed mathematics tutors for students with learning disabilities, and for girls. She holds an Ed.D. in mathematics and science education and M.S. and B.S. degrees in computer science. She is the author of over 100 research articles, is a Fulbright Fellow, was an elected member of the executive committee of AIED, and received several best paper awards (AIED 2009, 2010 and EDM 2010), one of them entitled "Emotion Sensors Go to School."

Dr. Jo-Anne Baird (GBR and SWE) is the Director of the Department of Education at the University of Oxford. Her research interests include examination standards, policy and systemic aspects of assessment, e-assessment and human judgment in assessment. She was the President of the Association for Educational Assessment-Europe from 2013 – 2015. Her first degree and doctorate were in psychology and she also has an MBA. She is actively involved in the development of many assessment policies and practices and serves as a member of PISA 2018 Reading Framework Advisory Group for the Global Competency Framework, Member of OECD Technical Review Panel, Guideline Reviewer for International Test Commission, Member of Pearson Assessment Expert Panel, DCSF Independent Advisor to the Expert Group, Member of the DCSF 14-19 Expert Group, Member of the Single Level Test Evaluation Group, National Assessment Agency Independent Advisor on national curriculum standard setting, Chair of the Ofqual Technical Advisory Group for their research program on examination reliability. She is now an Executive Editor of the journal Assessment in Education: Principles, policy and practice.

Dr. Ryan Baker (USA) is an associate professor in the Graduate School of Education at the University of Pennsylvania. His primary appointment is in the Teaching, Learning, and Leadership Division. He directs the Penn Center for Learning Analytics. He also has an affiliate appointment at Worcester Polytechnic Institute, in the Department of Social Science and Policy Studies, and courtesy appointments in the Department of Human Development at Teachers College Columbia University and at the University of Edinburgh Moray House School of Education. At Columbia University he was the Julius and Rosa Sachs Distinguished Lecturer at Teachers College, Columbia University. He earned his Ph.D. in human-computer interaction from Carnegie Mellon

University. Baker was previously Assistant Professor of Psychology and the Learning Sciences at Worcester Polytechnic Institute, and he served as the first technical director of the Pittsburgh Science of Learning Center DataShop, the largest public repository for data on the interaction between learners and educational software. He is currently serving as the founding president of the International Educational Data Mining Society, and as associate editor of the *Journal of Educational Data Mining*. His research combines educational data mining and quantitative field observation methods in order to better understand how students respond to educational software, and how these responses impact their learning. He studies these issues within intelligent tutors, simulations, multi-user virtual environments, and educational games, within populations from pre-schoolers, to middle school students, to military trainees.

Dr. Tiffany Barnes (USA) is an associate professor of computer science at North Carolina State University, who leads research in using data to personalize learning experiences, in creating games for education, exercise and energy, and in broadening participation. She received an NSF Faculty Early Career Development (CAREER) Award for her novel work in using data to add intelligence to Science, Technology, Engineering, and Mathematics (STEM) learning environments, and an NSF Improving Undergraduate STEM Education (IUSE) award to combine data-driven hints with data-driven pedagogical choices for learning in logic, probability, and programming. She co-leads the NSF Students and Technology in Academia, Research, and Service (STARS) Computing Corps that engages college students and faculty in tiered mentoring and leadership in outreach, research and service, and the NSF Beauty and Joy of Computing (BJC)-STARS project to develop faculty and teacher leaders to scale professional development for the new Computer Science Principles course. She serves on executive boards for the Association for Computing Machinery (ACM) Special Interest Group on Computer Science Education (SIGCSE), the International Educational Data Mining Society, the Artificial Intelligence in Education Society, and the IEEE Special Technical Community on Broadening Participation (STCBP). She has been on the organizing committees for several conferences, including SIGCSE, Educational Data Mining, and the Foundations of Digital Games. In 2015, she founded the IEEE STCBP's annual conference on Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT).

Dr. Chris Berka (USA) is the CEO and Co-Founder at Advanced Brain Monitoring (ABM). She has over 25 years' experience managing clinical research and developing and commercializing new technologies, is co-inventor of eleven patented and 14 patent-pending technologies, and served as the principal investigator or co-investigator for grants and contracts awarded by the National Institutes of Health, DARPA, ONR and NSF that provided more than $30 million of research funds to ABM. Her research in linking neuroscience and team processes is noteworthy for its potential contributions to the development of adaptive team tutors. Prior to founding ABM, Ms. Berka played a key role in the growth of an AMEX public company that patented and commercialized a forensic hair test for psychopharmacological profiles of drug use. She has 10 years' experience as a research scientist with publications on the analysis of the EEG correlates of cognition in healthy subjects and patients with sleep and neurological disorders. She received her BA with distinction in Psychology/Biology at Ohio State University and completed graduate studies in Neuroscience at UCSD.

Dr. Gautam Biswas (USA) is a Professor of Computer Science, Computer Engineering, and Engineering Management in the EECS Department and a Senior Research Scientist at the Institute for Software Integrated Systems at Vanderbilt University. He conducts research in Intelligent Systems with primary interests in analysis of complex embedded systems, and simulation-based environments for learning and instruction. This includes the Teachable Agents project, where students learn science by building causal models of natural processes and CTSiM, which exploits the synergy between computational thinking ideas and STEM learning to develop systems that help students learn science and math concepts by building simulation models. He has developed innovative educational data mining techniques for studying students' learning behaviors and linking them to metacognitive strategies. His research has been supported by funding from NASA, NSF, DARPA, and the US Department of Education. He has published extensively, and has over 400 refereed publications. Dr. Biswas is an associate editor of the *IEEE Transactions on Systems, Man, and Cybernetics*, *IEEE Transactions on Learning Technologies*, and the *Metacognition and Learning* journal. He is a Fellow of the IEEE, and member of the ACM, AAAI, and the Sigma Xi Research Society.

Dr. Stephen Blessing (USA) is currently an Associate Professor of Psychology at the University of Tampa. He has over 20 years of experience in the field of ITSs, starting with developing the Demonstr8 authoring tool while an intern at Apple Computer. He worked for 5 years at Carnegie Learning, creating the cognitive models for their high school math tutors. While there he started work on their Cognitive Tutor Software Development Kit, which allowed for the rapid creation of their model-tracing tutors. Dr. Blessing has maintained his research interest in authoring tools for ITSs, co-editing a book on the topic. He has collaborated with Dr. Stephen Gilbert on the Extensible Problem-Solving Tutor (xPST), another authoring tool that aims to make tutor creation easier and more affordable. He has an interest not only in how ITSs are used in traditional classroom environments, but also how they may be used in informal learning environments as well. He is currently testing an iPad-based tutor in a children's museum.

Dr. Amy Bolton (USA) is a Program Officer at ONR. She manages several programs within the Capable Manpower Future Naval Capability. Capable Manpower is a multi-million dollar per year science and technology (S&T) program that addresses the human system integration topics of manpower, personnel, training, and human system design. Products from Dr. Bolton's programs have had success transitioning to both the Navy and Marine Corps contributing to enhanced warfighter readiness across the Naval Enterprise. Dr. Bolton's research interests include adaptive training, human behavior modeling, human system design, and Live, Virtual, and Constructive training. She holds a PhD in applied experimental and human factors psychology from the University of Central Florida. Dr. Bolton has published more than 50 technical publications including four invited book chapters. Publications were on the topics of computational modeling, training technology and methodology, cognitive performance and resilience to stress, augmented cognition, and transitioning S&T into the acquisition process.

Dr. Keith Brawner (USA) is an adaptive training scientist at the US Army Research Laboratory in Orlando, Florida and a researcher in the Learning in Intelligent Tutoring Environments (LITE) Lab. He has 10 years of experience within US Army and Navy acquisition, development, and research agencies. He holds a Master's and PhD degree in computer engineering with a focus on intelligent systems and machine learning from UCF. His current research is in machine learning, active and semi-supervised learning, ITS architectures, and cognitive architectures. He manages research in adaptive training, semi/fully automated user tools for adaptive training content, and architectural programs toward next-generation training.

Dr. Bert Bredeweg (NLD) is an associate professor at the Informatics Institute within the University of Amsterdam (The Netherlands), leading the Qualitative Reasoning (QR) group. His research focus is the development of tools and expertise that support the acquisition of conceptual understanding of dynamic systems through conceptual modeling and simulation. Topics of interest include knowledge capture, QR, learning by modeling, cognitive diagnose, and HCI. He has supervised over 70 MSc and PhD students. He is an active member of the ecological informatics, QR, and educational technology communities. He is a regular senior reviewer and board member for the associated conferences and journals. He was part of the Computing Community Consortium (CCC)/NSF (USA) roadmap development (B. Woolf (Ed.), 2010), and consultant for SRI International (HALO project, 2010/2011), both USA, and acted four times as invited special issue editor for leading journals (most recently, guest editor for *IEEE Transactions on Learning Technologies* – special issue:

6(3), 2013, pp. 194–257 – together with Dr. B.M. McLaren and Dr. G. Biswas). He acquired and coordinated the DynaLearn project (EU FP7).

Dr. Peter Brusilovsky (USA) has been working in the area of adaptive systems and e-learning for many years. Since 1993 he has participated in the development of several adaptive web-based educational systems including ELM-ART, a winner of 1998 European Academic Software Award. He was involved in developing practical e-learning courses and systems as a Director of Computer Managed Instruction at Carnegie Technology Education, one of the first e-learning companies in the United States. Currently, he continues his research on adaptive e-learning as a professor of information science and intelligent systems at the University of Pittsburgh. He has published numerous research papers and several books adaptive systems and e-learning. He is the editor-in-chief of *IEEE Transactions on Learning Technologies* and a board member of several other journals. He is also the immediate past president of User Modeling, Inc., a professional association of user modeling researchers.

Dr. C. Shawn Burke (USA) is an Associate Professor (Research) at the Institute for Simulation and Training of the University of Central Florida. Her expertise includes teams and their leadership, team adaptability, team training, measurement, evaluation, and team effectiveness. Dr. Burke has published over 80 journal articles and book chapters and has presented/had work accepted at over 170 peer-reviewed conferences. She is currently investigating issues surrounding: leadership within virtually, distributed teams, team cohesion, issues related to multi-cultural team performance and multi-team systems. She is also currently funded on three separate NASA grants to investigate issues pertaining to:

1) Multi-cultural teams;
2) Team leadership; and
3) Social and task roles within long duration, exploration teams, respectively.

The above work is conducted with an interest in team leadership and training teams for operating in complex environments. Dr. Burke earned her doctorate in Industrial/Organizational Psychology from George Mason University and is an Associate Editor for the Journal of Trust Research and Consulting Editor for the Journal of Business and Psychology. She also serves as an ad-hoc reviewer for several journals, including: Leadership Quarterly, Journal of Applied Psychology, Military Psychology, Small Group Research. She has co-edited books on adaptability and advances in team effectiveness research. Her work in team modeling is on the critical path to realizing a collective tutoring capability in the future.

Dr. Winslow Burleson (USA) is an Associate Professor at New York University. He is the Founding Director of the Motivational Environments research group and author of over 100 scientific publications (including the "best paper" at AI in Ed 2009 and the 2011 UMUAI James Chen Award) and has been awarded 10 patents. In 2013, he was honored with a Google Faculty Research Award and in 2009 the National Academy of Engineering recognized him as, "One of the nation's brightest young engineering researchers and educators." He received his PhD from the MIT Media Lab, is a Kavli Fellow and serves on National Academy of Engineering, National Academies of Sciences, and National Science Foundation committees. Burleson has a BA in Biophysics from Rice University and an MSE in Mechanical Engineering from Stanford University.

Dr. Tak-Wai Chan (TWN) is Chair Professor of the Graduate Institute of Network Learning Technology at National Central University in Taiwan. He has worked on various areas of digital technology supported learning, including artificial intelligence in education, computer supported collaborative learning, digital classrooms, online learning communities, mobile and ubiquitous learning, digital game based learning, and, most recently, technology supported mathematics and language arts learning. In 1988, Chan produced his doctoral thesis, a seminal work in artificial intelligence in education, proposing the concept of virtual learning companion – a learning environment consisting of multiple virtual characters, such as a learning companion and a tutor. An environment with learning companions can provide a rich social context, allowing students to interact in diverse social activities including collaborative learning. The field of learning companion research is currently an active sub-area of artificial intelligence in education.

Dr. William J. Clancey (USA) is a senior research scientist at the Florida Institute for Human and Machine Cognition. His research relates cognitive and social science in the study of work practices and the design of agent systems. He has developed AI applications for medicine, finance, education, robotics, and spaceflight systems. He received a PhD in computer science at Stanford University and BA in mathematical sciences at Rice University. He was a founding member of the Institute for Research on Learning (1987 – 1997) and Chief Scientist of Human-Centered Computing at NASA Ames Research Center (1998 – 2013). He is a Fellow of the Association for Psychological Science, Association for Advancement of AI, National Academy of Inventors, and the American College of Medical Informatics. His seven books include *Working on Mars: Voyages of Scientific Discovery with the Mars Exploration Rovers* (recipient of AIAA 2014 Gardner-Lasser Aerospace History Literature Award). He has presented invited lectures in over 20 countries.

Dr. Joseph Cohn (USA) is a Commander in the US Navy's Aerospace Experimental Psychologist (AEP) community. He is currently assigned as Deputy Director to the Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics' (OUSD (AT&L)) Human Performance Training and BioSystems Directorate, with oversight for both the DoD's Human Research Protection Programs and the DoD's Human Systems and Medical research portfolios. He has previously served in the ONR's Human and Bioengineered Systems Division, as a Military Deputy and Program Officer and was ONR's first Deputy Director of Research for Science, Technology, Engineering and Mathematics (DDoR-STEM). Dr. Cohn also served as a program manager at the Defense Advanced Research Projects Agency (DARPA) directing basic and applied research projects that delivered cutting edge biomedical and information technology products, including deployable brain-imaging technologies, advanced brain-system interfaces, technologies that inoculate warfighters against stress, and a digital tutoring system that reduced by an order of magnitude the time required to train novices to perform at the expert level. He has co-authored over 80 publications, chaired numerous panels and workshops and been an invited speaker to national and international conferences on human systems research. He has co-edited a three-volume book series focusing on all aspects of training system development, and a single-volume book on enhancing human performance in high risk environments and is working on a book entitled *Modeling Sociocultural Influences on Decision Making*. He is a Fellow of the American Psychological Association, the Society of Military Psychologists, and Associate Fellow of the Aerospace Medical Association.

Dr. Ron Cole (USA) is Principal Scientist and President of Boulder Language Technologies, Inc. He received a BA in psychology from the University of Rochester and an MA and PhD in psychology from the University of California at Riverside. He established the Center for Spoken Language Understanding (CSLU) at the Oregon Graduate Institute, where he envisioned and managed development of the CSLU toolkit, with over 32,000 installations in 136 countries. The CSLU Toolkit was used as the research, development, and runtime platform to teach vocabulary to profoundly deaf children through spoken dialogue interaction with an animated computer character. The results of this multidisciplinary research effort were featured on ABC TV's Prime Time and the NSF Home Page on 03/2001 – 04/2001. He also co-founded the Center for Spoken Language Research at University of Colorado and established three successful companies. He has been principal investigator or co-PI on over $40 million in individual investigator peer-reviewed grants from NSF, National Institutes of Health (NIH), and Department of Education. His research with Wayne Ward at Boulder Language Technologies led to development of My Science Tutor, an ITS in which children learn spoken dialogues with a virtual tutor, with learning gains equivalent to expert human tutors. During the past 20 years, he has worked diligently to stimulate and sustain international collaboration in computer science and engineering. In 1997, with Jose Fortes, he organized the NSF-sponsored Workshop on International Collaboration in Computer Science. Subsequently, he organized several NSF-sponsored workshops with Jose Fortes, Jaime Carbonnell, and others in the US, Argentina, Chile, and Mexico to promote international collaboration in computer science. Several of these workshops led to new projects, initiatives and programs. He also managed a 2-year project sponsored by the NSF and EU to survey the state of the art of the field of human language technology, which resulted in an edited volume with contributions from over 90 authors.

Dr. Cristina Conati (CAN) is an associate professor of computer science at the University of British Columbia, Vancouver, Canada. She received a "Laurea" degree (M.Sc. equivalent) in computer science at the University of Milan, Italy (1988), as well as a M.Sc. (1996) and Ph.D. (1999) in intelligent systems at the University of Pittsburgh. Dr. Conati's research goal is to integrate research in artificial intelligence, cognitive science, and human-computer interaction to make complex interactive systems increasingly more effective and adaptive to the users' needs. Her areas of interest include intelligent user interfaces, user modeling, user-adaptive systems, and affective computing. Her research has received awards from the International Conference on User Modeling, the International Conference of AI in Education, the International Conference on Intelligent User Interfaces (2007), and the *Journal of User Modeling and User Adapted Interaction* (2002). Dr. Conati is an associate editor for the *Journal of AI in Education*, for the *IEEE Transactions on Affective Computing*, and the *ACM Transactions on Intelligent Interactive Systems*.

Dr. Nancy J. Cooke (USA) is a professor of Human Systems Engineering at Arizona State University and is Science Director of the Cognitive Engineering Research Institute in Mesa, AZ. She is also President-Elect of the Human Factors and Ergonomics Society, a member and chair of the National Research Council's Board on Human Systems Integration, and a chair of the NRC consensus study committee on The Science of Team Science. Dr. Cooke has organized annual workshops on the Human Factors of Unmanned Aerial Vehicles since 2004, has co-edited Human Factors of Remotely Operated Vehicles, published by Elsevier, The Best of Human Factors (with Eduardo Salas and published by HFES) and has co-authored (with Frank Durso), Stories of Modern Technology Failures and Cognitive Engineering Successes, published by Taylor and Francis.

Dr. Mark G. Core (USA) is a research scientist at the Institute for Creative Technologies (ICT) at the University of Southern California. He specializes in Artificial Intelligence (AI) in education working in ill-defined domains such as negotiation, cultural awareness, and leadership. He received his PhD from the University of Rochester in 2000 under the direction of Prof. Lenhart Schubert, and was a research fellow at the University of Edinburgh working with Prof. Johanna Moore until joining ICT in 2004. He worked in the area of computational linguistics specifically natural language understanding, and discourse analysis. At the University of Edinburgh, he undertook analysis of successful human tutoring dialogues and while at ICT he has focused on analysis of learner writing. At ICT, he is also working on an evolving tutoring architecture incorporating technologies such as expert modeling, open learner modeling, experience manipulation, explainable AI, natural language understanding, and natural language generation. Recent areas of research include authoring tools for tutoring systems, and physician training including interview and diagnosis skills.

Dr. Chris Dede (USA) is the Timothy E. Wirth Professor in Learning Technologies at Harvard's Graduate School of Education. His fields of scholarship include emerging technologies, policy, and leadership. His funded research includes seven current grants from NSF, Qualcomm, the Gates Foundation, and the US Department of Education Institute of Education Sciences to explore immersive simulations and transformed social interactions as means of student engagement, learning, and assessment. In 2007, he was honored by Harvard University as an outstanding teacher, and in 2011 he was named a Fellow of the American Educational Research Association. Dr. Dede has served as a member of the National Academy of Sciences Committee on Foundations of

Educational and Psychological Assessment and a member of the 2010 National Educational Technology Plan Technical Working Group.

Dr. Michel Desmarais (CAN) is professor at the Computer and Software Engineering Department of École Polytechnique de Montreal since 2002. His field of expertise is in the domains of HCI, e-learning, and AI. He has 15 years of experience in software project management. He was principal researcher of the HCI and computerized learning environments groups at the Computer Research Institute of Montreal between 1990 and 1998, where he directed a research program in HCI and computer assisted learning, and was involved in a number of research projects in close collaboration with private corporations. From 1998 to 2002, he was director of the web services department in a private company (www.mvm.com) and leader of a number of R&D web-based software projects. He is the editor of the *Journal of Educational Data Mining* and has authored over 100 scientific publications. His research interests include student and user modeling, user-centered software engineering, HCI, probabilistic modeling, and recommender interfaces.

Dr. Pierre Dillenbourg (CHE) is a full professor in learning technologies at École Polytechnique Fédérale de Lausanne (EPFL) in Switzerland. A former teacher in elementary school, Dr. Dillenbourg graduated in educational science (University of Mons, Belgium). He started his research on learning technologies in 1984. He obtained a PhD in computer science from the University of Lancaster (UK), in the domain of artificial intelligence applications for educational software. He has been assistant professor at the University of Geneva. He joined EPFL in 2002. As a professor in learning technologies in the School of Computer & Communication Sciences, he is the head of the CHILI Lab: "Computer-Human Interaction for Learning & Instruction". He is also the academic director of Center for Digital Education, which implements the MOOC strategy of EPFL. EPFL recently passed over 1 million MOOC registrations. Dr. Dillenbourg recently wrote a book entitled "Orchestration Graphs" that proposes a formal language for instructional design (EPFL Press).

Dr. Sidney D'Mello (USA) is an Assistant Professor in the departments of Computer Science and Psychology at the University of Notre Dame. His primary research interests are in the affective, cognitive, and learning sciences. More specific interests include affective computing, artificial intelligence in education, human-computer interaction, natural language understanding, and computational models of human cognition. He has co-edited five books and has published over 150 journal papers, book chapters, and conference proceedings in these areas. D'Mello's work on intelligent learning technologies including Affective AutoTutor, GazeTutor, ConfusionTutor, and GuruTutor has received seven outstanding paper awards at international conferences and has been featured in several media outlets including the *Wall Street Journal*. D'Mello serves on the executive board of the *International Artificial Intelligence in Education Society*, is a senior reviewer for the *Journal of Educational Psychology*, is an associate editor for *IEEE Transactions on Affective Computing*, and for *IEEE Transactions on Learning Technologies*. D'Mello received his PhD. in Computer Science from the University of Memphis in 2009. He also holds a M.S. in Mathematical Sciences and a B.S. in Electrical Engineering.

Dr. Michael Dorneich (USA) has research interests which focus on creating joint human-machine systems that enable people to be effective in the complex and often stressful environments found in aviation, military, robotic, and space applications. He specializes in adaptive systems which can provide assistance tailored to the user's current cognitive state, situation and environment. Adaptive systems are becoming more necessary as intelligent assistants are spreading into every aspect of work, education, and home life. His recent work explores the dual costs and benefits of adaptive systems, and develops approaches to mitigate shortcomings, leverage human strengths, and augment human performance when human capacity falls short of the demands of complex operational environments. Specific research and application areas include human factors, cognitive engineering, adaptive automation and adaptive interfaces, distributed systems, interactive learning environments, and decision-support systems. Prior to joining the faculty of IMSE, he worked at Honeywell Laboratories from

1999 – 2012 researching adaptive system design and human factors in a variety of domains. He holds 12 patents and has filed 20 additional. He has authored over 85 professional, peer-reviewed papers, and is currently an Associate Editor for the Journal of IEEE Transactions of Human-Machine Systems.

Dr. Benedict du Boulay (GBR) is an Emeritus Professor of Artificial Intelligence at the University of Sussex in Brighton. He was previously Dean (2002-2009) of the School of Science and Technology. He is currently President of the International Society for Artificial Intelligence in Education and is a member of the University of Sussex Department of Informatics, the Human Centered Technology Research Group and the Cognitive Sciences Research Center working in the areas of Artificial Intelligence in Education and the Psychology of Programming. He is also serving as an Erskine Visiting Fellow at the University of Canterbury, Christchurch, New Zealand.

Dr. Paula Durlach (USA) received her Ph.D. in experimental psychology from Yale University in 1982, and subsequently held fellowship positions at the University of Pennsylvania and the University of Cambridge. From 1987 to 1994, she was an assistant professor of psychology at McMaster University and then went on to lead the exploratory consumer science team at Unilever Research Colworth Laboratory in the U. K. She returned to the U. S. in 2001 to join the U. S. Army Research Institute for the Behavioral and Social Sciences. From 2012 – 2015, she served as the Deputy Director of the Advanced Distributed Learning Initiative. She is currently on the staff of US Army Research Laboratory where she maintains interest in adaptive training. Dr. Durlach has received recognition for her work in experimental psychology and cognitive science at the Army Science Conference and from the Department of Army Research and Development. She is a Fellow of the Association for Psychological Science, and member of the Experimental Psychology Society, the Psychonomic Society, and the Society for Artificial Intelligence in Education. With Dr. Alan Lesgold, she co-edited the book, *Adaptive Technologies for Training and Education*, published in 2012, and has also published research in journals such as *International Journal of Artificial Intelligence in Education*, *Military Psychology*, *Computers in Human Behavior*, and *Human-Computer Interaction*.

Dr. Stephen Fiore (USA) is faculty with the University of Central Florida's Cognitive Sciences Program in the Department of Philosophy and Director of the Cognitive Sciences Laboratory at UCF's Institute for Simulation and Training. He earned his Ph.D. degree in Cognitive Psychology from the University of Pittsburgh, Learning Research and Development Center. He maintains a multidisciplinary research interest that incorporates aspects of the cognitive, social, and computational sciences in the investigation of learning and performance in individuals and teams. He is co-Editor of recent volumes on Macrocognition in Teams (2008), Distributed Learning (2007), Team Cognition (2004), and he has co-authored over 100 scholarly publications in the area of learning, memory, and problem solving at the individual and the group level. As Principal Investigator and Co-Principal Investigator he has helped to secure and manage approximately $15 Million in research funding from organizations such as the National Science Foundation, the Office of Naval Research, the Air Force Office of Scientific Research, and the Department of Homeland Security.

Dr. J. Dexter Fletcher (USA) is a member of the senior research staff at the Institute for Defense Analyses where he specializes in personnel and human performance issues. His graduate degrees are in computer science and educational psychology from Stanford University, where, as a research associate, he directed projects for the Institute for Mathematical Studies in the Social Sciences. He has held university positions in psychology, computer science, and systems engineering and government positions in Navy and Army Service Laboratories, the Defense Advanced Research Projects Agency, and the White House Office of Science and Technology Policy. He has served on science and technology advisory panels for the Defense Science Board, Army Science Board, Naval Studies Board, Air Force Scientific Advisory Board, National Science Foundation, National Academy of Sciences, and the National Academy of Engineering. He has designed computer-based instruction programs used in public schools and training devices used in military training. He is a Fellow of the American Educational Research Association and three divisions of the American Psychological Association. His research interests include intelligent tutoring systems, synthetic environments in education and training, mobile performance aids, analyses of skilled behavior, and cost-effectiveness analyses of education and training.

Dr. Peter Foltz (USA) is a Vice President for research and development and works to bring innovative technologies to learning and assessment. He is one of the original developers of automated scoring technologies and holds a patent on methods for scoring of writing. Dr. Foltz's research has focused on language comprehension, 21st Century skills learning and assessment, and uses of machine learning and natural language processing in educational technology. The methods he has pioneered improve student achievement, expand student access, and make learning materials more affordable. He has lead the framework development for a new assessment of collaborative problem solving for the Organisation of Economic Cooperation and Development's Programme for International Student Assessment (PISA) test that will be given in 2015. A former professor of psychology at New Mexico State University, he has authored more than 85 journal articles, book chapters, conference papers, and other publications. He previously worked at Bell Communications Research and the Learning Research and Development Center at the University of Pittsburgh. Dr. Foltz holds doctorate and master's degrees in Cognitive Psychology from the University of Colorado, Boulder, and a bachelor's degree from Lehigh University.

Dr. Kenneth D. Forbus (USA) is the Walter P. Murphy Professor of Computer Science and Professor of Education at Northwestern University. He received his degrees from the Massachusetts Institute of Technology (MIT) (PhD in 1984). His research interests include QR, analogical reasoning and learning, spatial reasoning, sketch understanding, natural language understanding, cognitive architecture, reasoning system design, intelligent educational software, and the use of AI in interactive entertainment. He is a Fellow of the Association for the Advancement of Artificial Intelligence, the Cognitive Science Society, and the Association for Computing Machinery. He has received the Humboldt Award and has served as chair of the Cognitive Science Society.

Dr. Jared Freeman (USA) is Chief Scientist at Aptima, Inc. and has extensive experience investigating problem solving by individuals and teams in real-world settings. From this research, Dr. Freeman and his colleagues develop decision aids, training systems, measures of performance and communications, and organizational designs that support mission leaders and their staffs. Dr. Freeman is the author of more than 125 articles in journals, proceedings, and books concerning these and related topics. As Chief Scientist, Dr. Freeman is responsible for aligning Aptima's scientific and technical activities with the company's strategic goals. Dr. Freeman holds a Ph.D. in Human Learning and Cognition from Columbia University and a M.A. in Educational Technology from Teachers College, Columbia University. He is a Contributing Editor to the journal Human Factors and an occasional editor to other journals.

Dr. Dragan Gasevic (GBR) is a Professor and Chair in Learning Analytics and Informatics in the Moray House School of Education and the School of Informatics at the University of Edinburgh since February 2015. Before the current post, he was the Canada Research Chair in Semantic and Learning Technologies and a Professor in the School of Computing and Information Systems at Athabasca University since 2007. Presently, he is the President of the Society for Learning Analytics Research, an Adjunct Professor in the School of Interactive Arts and Technology at Simon Fraser University, Adjunct Professor in the School of Education at the University of South Australia, a Research Scientist in the LINK Research Lab at the University of Texas, Arlington, and an Honorary Adjunct Professor in the Department of Human Development at Teachers College, Columbia University. A computer scientist by training and skills, Dragan considers himself a learning and information scientist developing computational methods that can shape next-generation learning and software technologies and advance our understanding of information seeking, sense-making, and self-regulated and social learning. Funded by granting agencies and industry in Canada and Europe, Dragan is a recipient of several best paper awards at the major international conferences in learning and software technology. The award-winning work of his team on the LOCO-Analytics software is considered one of the pioneering contributions in the growing area of learning analytics. Recently, he has founded ProSolo Technologies Inc. that develops a software solution for tracking, evaluating, and recognizing competences gained through self-directed learning and social interactions. Committed to the development of international research community, Dragan had the pleasure to serve as a founding program co-chair of the International Conference on Learning Analytics and Knowledge in 2011 and 2012 and the general chair in 2016. Currently serving as a founding editor of the Journal of Learning Analytics and a program co-chair of the Learning Analytics Summer Institute, Dragan is a (co-)author of numerous research papers and books and a frequent keynote speaker.

Dr. Libby Gerard (USA) is a Research Scientist in the University of California (UC), Berkeley Graduate School of Education. Her research examines how innovative learning technologies can capture student ideas and help teachers and principals use those ideas to make decisions about classroom instruction and assessment. Her recent projects explore the use of automated scoring of student written essays and student created drawings to provide guidance to students and support the teacher. She designs and leads teacher and principal professional development by using student assessment data to inform instructional customization and resource allocation. Prior to being a research scientist, she was a postdoctoral scholar researcher at UC Berkeley where she coordinated the Mentored and Online Professional Development in Science (MODELS) project, and was a fellow of the Technology Enhanced Learning in Science (TELS) Center at Mills College. She also taught preschool and elementary school in Oakland, CA, and in Alessandria, Italy. Her research is published in leading peer-reviewed journals including *Science*, *Review of Educational Research* and *Journal of Research in Science Teaching*.

Dr. Stephen B. Gilbert (USA) received a BSE from Princeton in 1992 and PhD from MIT in 1997. He has worked in commercial software development and run his own company. He is currently an assistant professor in the Industrial and Manufacturing Systems Engineering Department at Iowa State University, as well as Associate Director of ISU's Virtual Reality Application Center and its Graduate Program in Human Computer Interaction. His research interests focus on technology to advance cognition, including interface design, intelligent tutoring systems, and cognitive engineering. He is a member of IEEE and ACM. He is currently the Principal Investigator for US Army Research Laboratory projects focused on future training technologies in the areas of adaptive team tutoring and virtual environments.

Dr. Benjamin S. Goldberg (USA) is an adaptive training scientist at the US Army Research Laboratory in Orlando, Florida and a member of the Learning in Intelligent Tutoring Environments (LITE) Lab. He has been conducting research in the Modeling and Simulation community for the past eight years with a focus on adaptive learning in simulation-based environments and how to leverage Artificial Intelligence tools and methods to create personalized learning experiences. Currently, he is the LITE Lab's lead scientist on instructional management research within adaptive training environments and is a co-creator of the Generalized Intelligent Framework for Tutoring (GIFT). Dr. Goldberg is a Ph.D. graduate from the University of Central Florida in the program of Modeling & Simulation. His work has been published across several well-known conferences, with recent contributions to the Human Factors and Ergonomics Society (HFES), Artificial Intelligence in Education and Intelligent Tutoring Systems (ITS) proceedings. Dr. Goldberg has also recently contributed to the journal Computers in Human Behavior and to the Journal of Cognitive Technology.

Dr. Ilya Goldin (USA) is a Director of Data Science at 2U, Inc., where his role is to use learning sciences and data science to advance 2U technology for students and faculty. He has engaged in research on a variety of topics in adaptive and personalized learning, including mastery modeling, domain modeling and help-seeking in tutoring systems; personal epistemology in learning science; and peer assessment in open-ended learning activities. Prior to 2U, he was a research scientist in the Center for Digital Data, Analytics and Adaptive Learning at Pearson, and an IES Post-doctoral Training Program in Interdisciplinary Education (PostPIER) postdoctoral fellow at CMU. He holds a PhD in intelligent systems from the University of Pittsburgh. He serves as an Edmund W. Gordon Fellow, funded by MacArthur Foundation and Educational Testing Service.

Dr. Robert Goldstone (USA) is Chancellor's professor in the Psychological and Brain Sciences Department and Cognitive Science Program at Indiana University, where he has been a faculty member since 1991. He received a BA from Oberlin College in 1986 in cognitive science, a Master's from University of Illinois in 1989, and a PhD in psychology from University of Michigan in 1991. His research interests include concept learning and representation, perceptual learning, educational applications of cognitive science, decision making, collective behavior, and computational modeling of human cognition. His interests in education focus on learning and transfer in mathematics and science, computational models of learning, and the design of innovative learning technologies. He was awarded two American Psychological Association (APA) Young Investigator awards in 1995 for articles appearing in the *Journal of Experimental Psychology*, the 1996 Chase Memorial Award for Outstanding Young Researcher in Cognitive Science, a 1997 James McKeen Cattell Sabbatical Award, the 2000 APA Distinguished Scientific Award for Early Career Contribution to Psychology in the area of Cognition and Human Learning, and a 2004 Troland research award from the National Academy of Sciences. He was the executive editor of *Cognitive Science* from 2001 – 2005, associate editor of *Psychonomic Bulletin & Review* from 1998-2000, and associate editor of *Cognitive Psychology* and *Topics in Cognitive Science* from 2007 – 2013. He was elected as a fellow of the Society of Experimental Psychologists in 2004, a fellow of the Association for Psychological Science in 2007, and a fellow of the Cognitive Science Society in 2006. From 2006 to 2011 he was the director of the Indiana University Cognitive Science Program.

Dr. Gregory A. Goodwin (USA) is a senior research scientist at the US Army Research Laboratory in Orlando, FL. For the last decade, he has worked for the Army researching ways to improve training methods and technologies. The current focus of his research is in learner modeling and methods for effectiveness evaluation of adaptive instructional techniques. He holds a PhD in psychology from Binghamton University and an MA in psychology from Wake Forest University.

Dr. Jamie Gorman (USA) received his PhD in Psychology from New Mexico State University in 2006 and completed his post-doctoral fellowship at Arizona State University from 2007-2010. His research interests include a variety of topics in human factors and basic research in cognitive and motor performance: communication analysis; dynamical systems theory; human-systems integration; team neurophysiology; team cognition; team coordination; team situation awareness; and team training and assessment. Dr. Gorman has published over 30 refereed articles and book chapters and is most well-known for his applications of dynamical systems theory and methods to understanding team performance. His research on real-time measurement of team dynamics, applications of nonlinear dynamics to interpersonal coordination, and the relationship between team communication and team neurodynamics has been funded by ONR, NSF, and DARPA. Dr. Gorman is a member of the Human Factors and Ergonomics Society (HFES) and serves on the editorial board of the journal Human Factors. He also serves on the 2016 Human Factors Prize review panel for HFES and in 2011 he and his co-authors received the Jerome H. Ely award from HFES for the best paper published in the 2010 volume of Human Factors.

Dr. Arthur Graesser (USA) is a professor in the Department of Psychology and the Institute of Intelligent Systems (IIS) at the University of Memphis (UofM), as well as an Honorary Research Fellow at University of Oxford. He received his PhD in psychology from the University of California at San Diego. His primary research interests are in cognitive science, discourse processing, and the learning sciences. More specific interests include knowledge representation, question asking and answering, tutoring, text comprehension, inference generation, conversation, reading, education, memory, emotions, Artificial Intelligence (AI), computational linguistics, and Human-Computer Interaction (HCI). He served as editor of *Discourse Processes* (1996 – 2005) and is the current editor of the *Journal of Educational Psychology* (2009 –2014). His service in professional societies includes president of the Empirical Studies of Literature, Art, and Media (1989 – 1992), the Society for Text and Discourse (2007 – 2010), the International Society for Artificial Intelligence in Education (2007 – 2009), and the Federation of Associations in the Behavioral and Brain Sciences Foundation (2012 – 2013). In addition to publishing over 600 articles in journals, books, and conference proceedings, he has written 3 books and co-edited 16 books. He and his colleagues have designed, developed, and tested software in learning, language, and discourse technologies, including AutoTutor, AutoTutor-Lite, AutoMentor, ElectronixTutor, MetaTutor, GuruTutor, DeepTutor, HURA Advisor, SEEK Web Tutor, Personal Assistant for Lifelong Learning (PAL3), Operation ARIES!, iSTART, Writing-Pal, Point & Query, Question Understanding Aid (QUAID), QUEST & Coh-Metrix.

Dr. Jim Greer (CAN) is a Professor of Computer Science and Director of the University of Saskatchewan's Learning Centre, a unit responsible for faculty development, curriculum innovation, and academic student support. He has worked for over 25 years researching and developing advanced learning technologies to support students and instructors. His research more recently has turned to learning analytics, big data, and privacy preservation in learner profiling. The Learning Centre he directs offers programming ranging from early career faculty development, to transition programs for freshmen students, to curriculum change facilitation for academic programs, to professional skills-training for graduate students, to courses on teaching and the scholarship of teaching and learning.

Prof Dr Samuel Greiff (LUX) is research group leader, principal investigator, and ATTRACT-fellow at University of Luxembourg). He has been and continues to be involved in the 2012, 2015, and 2018 cycle of the Programme for International Student Assessment (PISA), for instance as external advisor to the PISA 2012 and 2015 Expert and Subject Matter Expert Groups and as contracting partner at his institution. In this, he has considerably shaped the understanding of problem solving in PISA 2012 and of collaborative problem solving in PISA 2015. He has been working for several years on the assessment of transversal skills such as complex and collaborative problem solving and their role in the classroom, at work, and in private life. Currently, he is involved in the large-scale assessment of problem solving, collaboration, and life-long learning in various populations and leads a team of test developers, research assistants and graduate students dedicated at increasing the understanding, the measurement, and the application of different aspects of transversal skills and lifelong learning in educational contexts.

Dr. Ani Grubišić (HRV) is an assistant professor in the Department of Computer Science at the University of Split in Croatia. Her research interests include ITS design, student modeling in ITSs, e-learning and e-learning systems, evaluating the educational influence of e-learning systems, and autonomous, dynamic and adaptive courseware tools and methods. Dr. Grubišić has several projects related to ITS science and design, and was an invited speaker at the NATO Human Factors and Medicine Panel's HFM-237 Research Task Group on the "Assessment of Intelligent Tutoring System Technologies and Opportunities" in Bonn, Germany in 2014.

Dr. Glenn Gunzelmann (USA) is a Senior Research Psychologist in the in the Air Force Research Laboratory's Human Effectiveness Directorate and the Science and Technology Advisor for the Cognitive Models and Agents Branch. His research is focused on developing technologies, grounded in cognitive science research, to support training, including software-based intelligent tutors to provide individualized instruction for trainees, synthetic teammates for training events, and cognitively valid readiness tracking systems. His background is in computational cognitive modeling and leads research in the area of spatial cognition and understanding the effects of fatigue on cognitive performance and behavior.

Dr. Jiangang Hao (USA) is a research scientist at the Computational Psychometrics Research Center at Educational Testing Service (ETS) in Princeton, New Jersey. His current research focuses on collaborative problem solving, game and simulation-based assessment, educational data mining and analytics, and automated scoring. He co-leads the infrastructure sub-initiative of the game, simulation and collaboration initiative at ETS, and am the principal investigator of several research projects at ETS for designing simulation-based assessments, web-based platform for collaborative assessments and data analytics packages for game-based assessments. Dr. Hao received his Ph.D. in Physics and MA in Statistics, both from the University of Michigan. Prior to joining in ETS, he worked on modeling and mining Terabyte-scale data in astrophysics at Fermi National Accelerator Laboratory. His work has been reported by leading technology websites, such as the Wired and MIT Technology Review. Dr. Hao is the author of over 40 peer reviewed papers, with over 2500 total citations and h-index of 26.

Dr. Ioannis Hatzilygeroudis (GRC) is an Associate Professor in the Department of Computer Engineering and Informatics at University of Patras, Greece. She is also the Artificial Intelligence Group Director where the focus of her research is on knowledge representation (KR) with an emphasis on integrated/hybrid KR Schemes. Her research interests also include: knowledge-based/expert systems, intelligent decision support systems for medicine, and intelligent education systems. One of her primary projects is the Artificial Intelligence Teaching System (AITS). AITS is an ITS that assists human tutors in teaching and students in learning about artificial intelligence topics. AITS utilizes semantic web technologies, such as ontologies and semantic rules, to model domain knowledge, the learner's profile and the pedagogical model (curriculum sequencing specification, exercise selection, metacognitive choices). Semantic rules facilitate AITS to deliver personalized learning activities and to achieve adaptation to the student's characteristics. It also combines interactivity and visualization for implementing active learning principles as well as provide systematic feedback during interaction, based on as systematic error analysis. Visualizations implement processes in a step-by-step interactive way and students are invited to follow it. Furthermore, it includes mechanisms for automatic creation and difficulty estimation of exercises. Additionally, it offers mechanisms for automatic assessment of students' answers to exercises. Finally, it includes a learning analytics unit that records, analyses and visualizes learning data about the students' activities and progress. At the moment, AITS deals with two major subjects of AI, search algorithms and logic as a knowledge representation language, its content is written in Greek language, has been used in our department since the academic year 2012-2013 and has demonstrated very encouraging results.

Dr. Ming Hou (CAN) is a senior defence scientist at Defence Research and Development Canada (DRDC) Toronto, where he is responsible for providing science-based advice to the Canadian Armed Forces about the investment in and application of advanced technologies for human-machine systems requirements. His research interests include applied cognition, intelligent adaptive interface and systems design, human-technology/ automation interaction, intelligent tutoring, and stereoscopic virtual and mixed reality displays. Dr. Hou is the Canadian National Leader of the Human Systems Performance Technical Panel for the Air in The Technical Cooperation Program (TTCP). He also serves several NATO working groups. Dr. Hou is a senior member of the Institute of Electrical and Electronics Engineers (IEEE), a member of the Human Factors and Ergonomics Society, and a member of the Association of Computing Machinery.

Dr. Neil Heffernan (USA) is a Professor of Computer Science and Co-Director of the Learning Science & Technologies Program at Worcester Polytechnic Institute (WPI). For his dissertation from CMU, he built the first ITS that incorporated a model of tutorial dialogue. This system was shown to lead to higher student learning, by getting students to think more deeply about problems. It is based upon detailed studies of students, which produced basic cognitive science research results on the nature of human thinking and learning. He has written over 60 strictly peer-reviewed publications, and received multiple awards from professional associations. Since coming to WPI, he has received over a dozen major grants from NSF including the prestigious CAREER award, the US Department of Education, ONR, the US Army, the Massachusetts Technology Transfer Center, the Bill and Melinda Gates Foundation, and the Spencer Foundation worth over 13 million dollars. Recently, his work was cited in the National Educational Technology Plan and featured in the *NY Times Sunday Magazine*. Dr. Heffernan is best known for his role in the development of the ASSISTments service, which helps students learn mathematics even as it assesses their knowledge, and which is used by over 50000 students a year in the US. He is widely published in intelligent tutoring systems, and educational data mining. His work gained prominence when a New York Times Magazine story by Annie Murphy Paul featured ASSISTments and Heffernan's research with the tool.

Dr. Xiangen Hu (USA and CHN) is a professor in the Department of Psychology and Department of Electrical and Computer Engineering at the University of Memphis (UofM) and senior researcher at the Institute for Intelligent Systems (IIS), and a visiting professor at Central China Normal University (CCNU). He received his MS in applied mathematics from Huazhong University of Science and Technology, MA in social sciences, and PhD in cognitive sciences from the University of California, Irvine. He is the Director of Advanced Distributed Learning (ADL) Center for Intelligent Tutoring Systems (ITSs) Research & Development and a senior researcher in the Chinese Ministry of Education's Key Laboratory of Adolescent Cyberpsychology and Behavior. His primary research areas include mathematical psychology, research design and statistics, and cognitive psychology. More specific research interests include General Processing Tree (GPT) models, categorical data analysis, knowledge representation, computerized tutoring, and advanced distributed learning. He receives funding for the above research from the US National Science Foundation (NSF), US Institute for Education Sciences (IES), ADL of the US Department of Defense (DOD), US Army Medical Research Acquisition Activity, US Army Research Laboratories (ARL), US Office of Naval Research (ONR), UofM, and CCNU.

Dr. Ronghuai Huang (CHN) is a Professor in Education and Dean of the Smart Learning Institute at Beijing Normal University (BNU) in which educational technology is one of the National Key Subjects. He is also director of R&D at the Center for Knowledge Engineering, which is dedicated to syncretizing artificial intelligence and human learning. Prof. Huang has being engaged in research on educational technology as well as knowledge engineering since 1997. He has accomplished or is working on over 60 projects, including those key science and technology projects to be tackled in the national "Ninth Five-year Plan", "Tenth Five-year Plan" and "Eleventh Five-year Plan" and the projects in the national 863 plan as well as others financed by the Chinese government. His ideas have been widely spread, with more than 180 academic papers and over 20 books published both domestically and overseas.

Dr. Seiji Isotani (BRA) is an Educational Technologist, Scientist and Innovator. He holds the positions of Associate Professor of Computer Science and Co-Director of the Applied Computing in Education Laboratory at the University of São Paulo. He is also the co-founder of two start-up companies that have won several innovation awards in the field of education and semantic technology. His main research topics are: Intelligent Tutoring Systems (ITS), Computer-Supported Collaborative Learning (CSCL), Ontologies and Semantic Web, Dynamic/Interactive Geometry, and technology-enhanced learning. Isotani's research group focuses on understanding how computational technologies can be designed and improved to create smart learning environments that help students to achieve robust learning.

Dr. G. Tanner Jackson (USA) is a research scientist in the Research and Development Division at Educational Testing Service (ETS) in Princeton, NJ. Tanner received a PhD degree in cognitive psychology in 2007 and a MS degree in cognitive psychology in 2004 – both from the University of Memphis. He also received a BA degree in psychology from Rhodes College in 2001. After completing a Postdoctoral Fellowship at UofM (2008 – 2011), he continued his research as an Assistant Research Professor within the Learning Sciences Institute at ASU (2011-2013). His current work at ETS focuses on innovative assessments and student process data. His main efforts involve the development and evaluation of conversation-based formative assessments (through ETS strategic initiatives) and game-based assessments (working in collaboration with GlassLab). Additionally, he is interested in how users interact with complex systems, and leverages these environments to examine and interpret continuous and live data streams, including user interactions across time within an assessment system.

Dr. W. Lewis Johnson (USA) co-founded Alelo in 2005 as a spin out of the University of Southern California. Under his leadership, Alelo has developed into a major producer of innovative learning products focusing on communication skills. Alelo has developed courses for use in a number of countries around the world, all using the Virtual Role-Play method. Dr. Johnson is an internationally recognized leader in innovation for education and training. In 2012, he was keynote speaker at the International Symposium on Automated Detection of Errors in Pronunciation Training in Stockholm. In 2013, he was keynote speaker at the International Association of Science and Technology for Development (IASTED) Technology Enhanced Learning Conference and the SimTecT conference, and was co-chair of the Industry and Innovation Track of the Artificial Intelligence in Education (AIED) 2013 conference. In 2014, he was keynote speaker at the International Conference on Intelligent Tutoring Systems, and was Distinguished Lecturer at the National Science Foundation. Recently, Dr. Johnson led the development of Adaptive Refresher Training (ART) which uses role-playing technology combined with adaptive learning to create customized refresher training for language skills. Language skills will decay over periods of time such as summer vacation or delays between training and usage. The ART project seeks to quickly recover language skills after attrition has occurred by identifying specific training needs and providing personalized instruction using simulated face-to-face conversations and language instruction activities. When not engaged in developing disruptive learning products, Lewis and his wife Kim produce Kona coffee in Hawaii.

Dr. Joan Johnston (USA) has been a US military research psychologist for 25 years. Her current research focus is on training effectiveness with an emphasis on training transfer. Dr. Johnston's areas of expertise include training and decision support systems for tactical decision making under stress, team performance and team training technologies, embedded and distributed simulation-based training, leadership and operational readiness in joint and multinational exercises, and cross-cultural competence. She recently joined the staff of the US Army Research Laboratory as a Senior Scientist. Prior to this she was the Army Research Institute Orlando Unit Chief, developing principles and guidelines for employing adaptive training technologies and mobile platform learning environments. She was a senior research psychologist for 22 years with the Naval Air Warfare Center Training Systems Division. Dr. Johnston earned an M.A. and a Ph.D. in I/O Psychology from the University of South Florida.

Dr. Irvin R. Katz (USA) is Director of the Cognitive Sciences Research Group at ETS in Princeton, NJ. He earned a PhD in cognitive psychology from CMU in 1988. In addition to ETS, he has held positions at Keio University in Yokohama, Japan; the US Bureau of Labor Statistics; the US Census Bureau; and George Mason University. Throughout his 25-year career at ETS, he has conducted research that applies and develops theories of cognitive learning and reasoning to issues of educational assessment. Dr. Katz is also a human-computer interaction practitioner with more than 30 years of experience in designing, building, and evaluating software for research, industry, and government. The Cognitive Science Research Group that he directs comprises

12 scientists who conduct research and development at the forefront of educational assessment, using cognitive theory in the design of assessments, building cognitive models to guide interpretation of test-takers' performance, and researching cognitive issues in the context of assessment. Moving beyond traditional (e.g., multiple-choice) tests, the group investigates reliable and valid assessment (both summative and formative) using innovative, highly interactive digital environments such as online games, virtual labs or other simulations, and human-agent conversation-based interactions.

Dr. Judy Kay (AUS) is Professor of Computer Science. She leads the Human Centered Technology Research Cluster, one of three priority clusters in the Faculty of Engineering and IT at the University of Sydney. Her own lab, CHAI, Computer Human Adapted Interaction Research Group aims to create new technologies for human computer interaction (HCI). Her personalization research has created the Personis user modeling framework. This is a unified mechanism for keeping and managing people's long term personal data from diverse sources. This is the foundation for building personalized systems. Personis models are distinctive in that they were designed to be scrutable, because interfaces enable the user to scrutinise their user model and personalization processes based on it. In learning contexts, she has created interfaces for Open Learner Models that make this personal data available in useful forms for long term learning and self-monitoring. Her interface research has created the Cruiser Natural User Interaction (NIU) software framework. This provides new ways for people to make use of large interactive tabletops and wall displays. By mining the digital footprints of such interaction, this research is creating new ways for people to learn to collaborate, and to learn and work more collaboratively. Dr. Kay is widely published.

Dr. Kenneth Koedinger (USA) is Professor of Human-Computer Interaction and Psychology at CMU. His research has contributed new principles and techniques for the design of educational software and has produced basic cognitive science research results on the nature of student thinking and learning. Dr. Koedinger is a co-founder of Carnegie Learning (http://carnegielearning.com) and the CMU Director of LearnLab (http://learnlab.org). LearnLab is supporting Big Data investigations in education and, more generally, leverages cognitive and computational approaches to support researchers in investigating the instructional conditions that cause robust student learning. See http://pact.cs.cmu.edu/koedinger.html for more information.

Dr. Rohit Kumar (USA) is a Senior Scientist in the Speech, Language and Multimedia division of Cambridge, Massachusetts based Raytheon BBN Technologies. His work aims to amplify human productivity and enrich the experience of interaction with machines through the use of intelligent technologies. His research trajectory is advancing the use of intelligent, yet principally imperfect technologies in a broad range of end-user facing applications. At BBN, Dr. Kumar led BBN's efforts on bringing together problem-solving, a time-tested pedagogical practice across numerous STEM disciplines, with recent advances in intelligent tutoring system. BBN Learning Platform, the result of this work, is now being made available, free of cost, to high schools across the nation. Rohit has now taken on the challenge of applying the underlying tutoring technologies developed for Physics, to advance the state of the art in simulation-based training for robot assisted surgery. In addition to his work on education and training technologies, Rohit has been the Technical Lead for BBN's advanced speech translation program that is focused on creating technological mediums for cross-lingual communication. His work has increased the resilience of these translation systems to component failures and made them rapidly adaptable for timely deployment in emergent scenarios like humanitarian aid and disaster relief. Dr. Kumar is a senior member of IEEE and has authored over 60 publications. He graduated from Punjab Engineering College, Chandigarh and received his PhD from Carnegie Mellon University, Pittsburgh. His dissertation work studied conversational agents that improve the effectiveness of human collaboration. This work demonstrated that the use of social interaction by conversationally capable machines (e.g., chat bots) can be used to manage user attention. Applied to educational scenarios, where the students are easily distracted, regulated social interaction helps in focusing the student's attention and improving learning outcomes.

Dr. Vivekanandan Kumar (CAN) is a full professor in the School of Computing and Information Systems at Athabasca University, Canada. He holds the Natural Sciences and Engineering Research Council of Canada's (NSERC) Discovery Grant on Anthropomorphic Pedagogical Agents, funded by the Government of Canada. His research focuses on developing anthropomorphic agents, which mimic and perfect human-like traits to better assist learners in their regulatory tasks. His research includes investigating technology-enhanced erudition methods that employ Big Data Learning Analytics, Self-Regulated Learning, Co-Regulated Learning, Causal Modeling, and Machine Learning, to facilitate deep learning. He earned his doctoral degree in Computer Science (PhD) from the University of Saskatchewan, Canada, as the best graduating student in 2001. He holds a master's degree in Computer Science Applications (MCA) as the first rank holder from Bharathiar University, India, where he also completed his bachelor's degree (BSc) in Physics with a minor in Mathematics.

Dr. Bor-Chen Kuo (TWN) received the B.S. and M.S. degrees in mathematics education and educational statistics, respectively, from National Taichung Teachers College, Taiwan, R.O.C., in 1993 and 1996, and the Ph.D. degree in electrical and computer engineering from the Purdue University, West Lafayette, IN, in 2001. He is currently a Distinguished Professor in the Graduate Institute of Educational Information and Measurement, and the Dean of College of Education, National Taichung University of Education, Taiwan. Dr. Kuo is the president of Chinese Association of Psychological Testing and the Chief Editor of the Journal of Educational Measurement and Statistics, Taiwan. He also serves as Guest Editors in many international journals and Editorial Board Member of Journal of Educational Measurement. Dr. Kuo received an Outstanding and Excellence Research Award from the R.O.C Education and Research Society in 2009. His research interests include computerized adaptive testing, cognitive diagnostic modeling, machine learning, and artificial intelligence in education.

Dr. Susanne Lajoie (CAN) is a Professor and Canadian Research Chair Tier 1 in Advanced Technologies for Learning in Authentic Settings in the Department of Educational and Counselling Psychology at McGill University and a member of the Centre for Medical Education. She is a Fellow of the American Psychological Association, appointed for her outstanding contributions to the field of Psychology as well as a Fellow of the American Educational Research Association. She received her Doctorate from Stanford University in 1986. Dr. Lajoie is a recipient of the McGill Carrie Derick Award for graduate supervision and teaching. Dr. Lajoie is the Director of the Learning Environments Across Disciplines partnership grant funded by the Social Sciences and Humanities Research Counsel in Canada. Her research involves the design of technology rich learning environments for educational and professional practices. She explores how theories of learning and affect can be used to guide the design of advanced technology rich learning environments in different domains, i.e., medicine, mathematics, history, etc. These environments serve as research platforms to study student engagement and problem solving in authentic settings. She uses a cognitive approach to identify learning trajectories that help novice learners become more skilled in specific areas and designs computer tools to enhance self-regulation, memory, and domain-specific learning. She has numerous publications and has been invited to present her research worldwide including Australia, France, Germany, Hong Kong, Korea, Singapore, Spain, Sweden, Taiwan, Mexico, the UK and the Ukraine.

Dr. Michelle LaMar (USA) is an Associate Research Scientist in the Computational Psychometrics Research Center at Educational Testing Service. Her current research focuses on the development of psychometric models appropriate for use with complex assessment tasks such as simulations or games. She is particularly interested in modeling task-process data using dynamic cognitive models to enable valid inference about multiple layers of student cognition. Michelle received a Masters in curriculum studies from Sonoma State University and a Ph.D. in educational measurement from the University of California, Berkeley. Prior to her doctoral work, Michelle spent 18 years in software engineering, specializing in educational simulations, authoring tools, and natural-language parsing.

Dr. David Landy (USA) is an assistant professor in the psychological and brain sciences at Indiana University in Bloomington, IN. He also earned his PhD in there in 2007 in computer science and cognitive science. In between, he was a post-doctoral research scientist at the University of Illinois at Urbana-Champaign and an assistant professor at the University of Richmond. His work on perceptual aspects of notational reasoning earned the 2007 David Marr prize from the Cognitive Science Society, and a new investigator award from the APA in 2008. His work on interactive and dynamic algebras has been funded by the Institute for Education Sciences.

Dr. H. Chad Lane (USA) is an associate professor of Educational Psychology and Informatics at the University of Illinois, Urbana-Champaign (UIUC). His work focuses on the application of AI and entertainment technologies to educational problems. He has published over 40 papers in areas including educational games, pedagogical agents, scaffolding/feedback, and virtual environments for learning. Prior to joining UIUC, he was the Director for Learning Sciences Research at the USC ICT. He received his PhD in Computer Science in 2004 from the University of Pittsburgh where he studied intelligent tutoring systems and the learning sciences. In 2013, Chad served as the Program Co-Chair for the 16th International Conference on AIED. He also serves on the executive committee of the AIED Society (an elected position), as an associate editor for several major educational technology journals, and as an advisor for the NSF Cyberlearning CIRCL center. More information is available on his website: http://hchadlane.net.

Dr. Walter S. Lasecki (USA) is an assistant professor of computer science and engineering at the University of Michigan, Ann Arbor, where he directs the Crowds+Machines (CROMA) Lab. He and his students create interactive intelligent systems that are robust enough to be used in real-world settings by combining both human and machine intelligence to exceed the capabilities of either. These systems let people be more productive, and improve access to the world for people with disabilities. He received his PhD and MS from the University of Rochester in 2015 and a BS in computer science and mathematics from Virginia Tech in 2010. He has previously held visiting research positions at CMU, Stanford, Microsoft Research, and Google[x].

Dr. Joseph J. LaViola Jr. (USA) is the Charles N. Millican Faculty Fellow and associate professor in the Department of Electrical Engineering and Computer Science and directs the Interactive Systems and User Experience Research Cluster of Excellence at UCF. He is the director of the M&S graduate program and is also an adjunct associate research professor in the Computer Science Department at Brown University. His primary research interests include pen-based interactive computing, 3D spatial interfaces for video games, human-robot interaction, multimodal interaction in virtual environments, and user interface evaluation. His work has appeared in journals such as *ACM TOCHI*, *IEEE PAMI*, *Presence*, and *IEEE Computer Graphics & Applications*, and he has presented research at conferences including ACM CHI, ACM IUI, IEEE Virtual Reality, and ACM SIGGRAPH. He has also co-authored *3D User Interfaces: Theory and Practice*, the first comprehensive book on 3D user interfaces. In 2009, he won an NSF Career Award to conduct research on mathematical sketching. He received a ScM in computer science in 2000, a ScM in applied mathematics in 2001, and a PhD in computer science in 2005 from Brown University.

Dr. Douglas B. Lenat (USA) is one of the world's leading computer scientists, and is the founder of the Cyc project and president of Cycorp. Dr. Lenat has been a Professor of Computer Science at CMU and Stanford University and has received numerous honors including awarded the biannual International Joint Conference on Artificial Intelligence (IJCAI) Computers and Thought Award, which the highest honor in AI; named first Fellow of the Association for the Advancement of Artificial Intelligence, and Fellow of the American Academy for the Advancement of Science. He is a prolific author, whose hundreds of publications include *Knowledge Based Systems in Artificial Intelligence* (1982, McGraw-Hill), *Building Expert Systems* (1983, Addison-Wesley), *Knowledge Representation* (1988, Addison-Wesley), and *Building Large Knowledge Based Systems* (1989, Addison-Wesley). His 1976 Stanford thesis earned him the biannual IJCAI Computers and Thought Award in

1977. He received his PhD in computer science from Stanford University and his BA and MS in mathematics from the University of Pennsylvania. He is an editor of the *J. Automated Reasoning*, *J. Learning Sciences*, and *J. Applied Ontology*. He is a founder and Advisory Board member of TTI Vanguard, and is the only individual to have served on the Scientific Advisory Boards of both Microsoft and Apple.

Dr. Alan Lesgold (USA) is professor and, since July 2000, dean of the School of Education at the University of Pittsburgh and also professor of psychology and intelligent systems. He received his Ph.D. in psychology from Stanford University in 1971 and also holds an honorary doctorate from the Open University of the Netherlands. He is a fellow of the American Psychological Association (APA), in experimental, applied, and educational psychology, and also of the Association for Psychological Science and the American Educational Research Association. In 2001, he received the APA award for distinguished contributions of applications of psychology to education and training. In 1995, he was awarded the Educom Medal. He was president of the Applied Cognitive Psychology division of the International Association for Applied Psychology 2002-2006. Lesgold is a Lifetime National Associate of the National Research Council (National Academies). He also was appointed by Governor Rendell as a member of the Governor's Commission on Preparing America's Teachers in 2005 and served later on the State's commission on cyber high schools as well. He served as chair of the National Research Council committee on adolescent and adult literacy. In that role, he led an extensive study of the systems, both in the K-12 world and in community colleges, for remediating literacy problems. From 1986 to 2000, he was executive associate director of the Learning Research and Development Center at the University of Pittsburgh. He is on the board of Teaching Matters. Lesgold has served the Pittsburgh community mostly in education-related activities, including board service for A+ Schools, Youthworks (completed), and the Pittsburgh Regional Center for Science Teaching. He serves as chair of the advisory board for the Center for Learning at Community College of Allegheny County. He was president of Rodef Shalom Congregation in 2002-2004. He is married to Sharon Lesgold, a retired mathematics educator, and they have two grown sons, Jacob and Noah.

Dr. James C. Lester (USA) is a Distinguished Professor of Computer Science and the Director of the Center for Educational Informatics at North Carolina State University. His research focuses on transforming education with technology-rich learning environments. Using AI, game technologies, and computational linguistics, he designs, develops, fields, and evaluates next-generation learning technologies for K-12 science, literacy, and computer science education. His work on personalized learning ranges from game-based learning environments and ITSs to affective computing, computational models of narrative, and natural language tutorial dialogue. The adaptive learning environments he and his colleagues develop have been used by thousands of students in K-12 classrooms throughout the US He received his BA (Highest Honors, Phi Beta Kappa), MSCS, and PhD in computer science from the University of Texas at Austin. He received his BA in history from Baylor University. He has served as Editor-in-Chief of the *International Journal of Artificial Intelligence in Education* and Program Chair for the International Conference on Intelligent Tutoring Systems, the International Conference on Intelligent User Interfaces, and the International Conference on Foundations of Digital Games. The recipient of a NSF CAREER Award, he is a Fellow of the Association for the Advancement of Artificial Intelligence (AAAI).

Dr. Marcia C. Linn (USA) is Professor of Development and Cognition, specializing in S&T in the Graduate School of Education, UC Berkeley. She is a member of the National Academy of Education and a Fellow of the American Association for the Advancement of Science (AAAS), the American Psychological Association, and the Association for Psychological Science. She has served as President of the International Society of the Learning Sciences, Chair of the AAAS Education Section, and on the boards of the AAAS, the Educational Testing Service Graduate Record Examination, the McDonnell Foundation Cognitive Studies in Education Practice, and the NSF Education and Human Resources Directorate. Awards include the National Association for Research in Science Teaching Award for Lifelong Distinguished Contributions to Science Education, the American Educational Research Association Willystine Goodsell Award, and the Council of Scientific Society Presidents first award for Excellence in Educational Research.

Dr. Noboru Matsuda (USA) is research faculty at the Human-Computer Interaction Institute at CMU. His primary research interest is in an application of cutting-edge technologies to build an effective learning technology for all students. To achieve this goal, Dr. Matsuda studies the transformative theory of advanced educational technology as well as cognitive theories of learning and teaching. Dr. Matsuda received a PhD in intelligent systems from the University of Pittsburgh in 2004. Dr. Matsuda has developed a number of ITSs in math (arithmetic, geometry theorem proving and algebra equations), C language, and the formal specification language Z. In recent years, Dr. Matsuda has been leading the SimStudent project (www.SimStudent.org) where the research team develops an AI that learns problem-solving skills through guided-problem solving (aka peer tutoring) and worked-out examples (aka learning by self-explanation). Applications of SimStudent include:

1) Developing an innovative authoring system for cognitive tutors by using SimStudent as an intelligent apprentice that learns subject matter knowledge from authors;
2) Understanding the theory of learn by teaching by using SimStudent as a synthetic peer that students can teach; and
3) Advancing theory of learning by running simulations using SimStudent.

Dr. Camillia Matuk (USA) is Assistant professor of Educational Communication and Technology at New York University's Steinhardt School of Culture, Education, and Human Development. Her interests are in the design of technologies for teaching, learning, and collaboration. Recently, she has been involved in researching how tools within online learning environments can support classroom science inquiry, and how they can encourage teachers to design and refine their instruction. Matuk has a PhD in learning sciences from Northwestern University, an MSc in biomedical communications from the University of Toronto, and a BSc in biological sciences from the University of Windsor. She completed a postdoctoral fellowship with the TELS center at UC Berkeley.

Dr. Danielle S. McNamara (USA) is a Professor in the Psychology Department at ASU and director of the Science of Learning and Educational Technology laboratory. She focuses on educational technologies and discovering new methods to improve students' ability to understand challenging text, learn new information, and convey their thoughts and ideas in writing. Her work integrates various approaches and methodologies including the development of game-based ITSs (e.g., iSTART, Writing Pal), the development of NLP tools (e.g., iSTART, Writing Pal, Coh-Metrix, the Writing Assessment Tool), basic research to better understand cognitive and motivational processes involved in comprehension and writing, and the use of learning analytics across multiple contexts. More information about her research and access to her publications are available at www.soletlab.com.

Dr. Antonija (Tanja) Mitrovic (NZL) is a full professor and the Head of the Department of Computer Science and Software Engineering at the University of Canterbury, Christchurch, New Zealand. She is the leader of Intelligent Computer Tutoring Group (ICTG). Dr. Mitrovic received her PhD in computer science from the University of Nis, Yugoslavia, in 1994. Prof. Mitrovic is president of the International Society of Artificial Intelligence in Education. She is an associate editor of the following journals: *International Journal on Artificial Intelligence in Education*, *IEEE Transactions on Teaching and Learning Technologies*, and *Research and Practice in Technology Enhanced Learning* (RPTEL). Dr. Mitrovic's primary research interests are in student modeling. ICTG has developed a number of constraint-based intelligent tutoring systems in a variety of domains, which have been thoroughly evaluated in real classrooms, and proven to be highly effective. These systems provide adaptive support for acquiring both problem-solving skills and meta-cognitive skills (such as self-explanation and self-assessment). Although most of the ITSs developed by ICTG support students learning individually in areas such as database querying (SQL-Tutor), database design (EER-Tutor and ERM-Tutor), and data normalization (NORMIT), there are also constraint-based tutors for object-oriented software design and collaborative skills, various engineering topics (thermodynamics, mechanics), training to interpret medical images and language-learning. ICTG has also developed the Authoring Software Platform for Intelligent Resources in Education (ASPIRE), a full authoring and deployment environment for constraint-based tutors. Recent research includes affect-aware tutors and motivational tutors. She has authored over 200 peer-reviewed publications.

Dr. Piotr Mitros (USA) is Chief Scientist at edX and its technical co-founder. Dr. Mitros is charged with developing and applying technology to optimize the learning process. As a graduate student, Mitros took breaks from his thesis to spend time teaching in China, working in India, and facilitating educational technology projects in Nigeria, as well as developing experimental educational formats at MIT. His observations of university systems around the world inspired Mitros to find innovative ways to dramatically increase both the quality of and access to education. Following a stint in industry designing the analog front end for a novel medical imaging modality at Rhythmia Medical, Dr. Mitros returned to MIT to lead the creation of the original MITx platform and help lead the creation of its pedagogy. He brings a broad interdisciplinary background that combines teaching, engineering, computer science, and math, and has been interested in teaching and education since he was a child. Mitros enjoys making things -- curtains, bicycle parts, electronics, furniture, and speakers. Dr. Mitros holds a B.S. degree in Math and Electrical Engineering, a Masters of Engineering in EECS, and a Ph.D. in EECS, all from MIT.

Dr. Johanna Moore (GBR) is Chair of Artificial Intelligence and Head of the School of Informatics at the University of Edinburgh. She previously held posts at UCLA (1976 – 1986), USC Information Sciences Institute (1983 – 1989), and the University of Pittsburgh (1990 – 1998). Since 1998 she has held the Chair in Artificial Intelligence, within the School of Informatics at the University of Edinburgh. She is also Director of the Human Communication Research Centre. Moore was elected to the UCLA chapter of Phi Beta Kappa in 1980. She held an Office of Naval Research Fellowship from 1982 to 1985, and was an International Business Machines Fellow from 1985 to 1987. She held a National Science Foundation National Young Investigator Award, from 1994 to 1999. She has been Chair of the Cognitive Science Society and was President of the Association for Computational Linguistics in 2004. She is currently a Fellow of the British Computer Society and of the Royal Society of Edinburgh. Her recent work on the Basic Electricity and Electronics Tutorial Learning Environment (BEETLE II) advanced the state of the art in dynamic adaptive feedback generation and natural language processing (NLP) by extending symbolic NLP techniques to support unrestricted student natural language input in the context of a dynamically changing simulation environment in a moderately complex domain.

Dr. Jason Moss (USA) is an adaptive training scientist at the US Army Research Laboratory in Orlando, Florida. He has over 14 years of experience conducting military-funded research and experimentation for the U.S. Navy and U.S. Army. He has successfully managed, conceptualized, and executed research programs in academia, industry, and the U.S. federal government. In addition to his research in adaptive training, Dr. Moss is a prominent researcher in the field of simulator sickness and Virtual Environments (VEs). Further, his body of research has focused on training and human-machine interactions with VEs, psychophysiological measures of workload, perceptual considerations related to VEs, training effectiveness of simulation and VE technologies, and training effectiveness of intelligent tutoring and adaptive training. Dr. Moss is published and has presented at numerous international conferences in the areas of simulator sickness, perception, and training effectiveness. Dr. Moss was a pioneering human factors graduate student as the first recipient of the Ph.D. in human factors psychology from Clemson University. During his graduate studies, Dr. Moss was the sole recipient of a national IITSEC (Interservice/Industry Training, Simulation and Education Conference) doctoral scholarship award and Clemson University's H.W. Close Distinguished Graduate Fellowship award.

Dr. Bradford Mott (USA) is a Senior Research Scientist in the Center for Educational Informatics at North Carolina State University. He received his PhD in computer science from North Carolina State University, where his research focused on intelligent game-based learning environments. His research interests include AI and human-computer interaction, with applications in educational technology. In particular, his research focuses on game-based learning environments, intelligent tutoring systems, computer games, and computational models of interactive narrative. His research has been recognized with best paper awards and he has contributed to several award-winning video games, including one that received a game of the year award. He has many years of software development experience from industry, including extensive experience in the video game industry, having served as Technical Director at Emergent Game Technologies where he created cross-platform middleware solutions for Microsoft's Xbox and Sony's PlayStation video game consoles.

Dr. Kasia Muldner (USA) received her Ph.D. from the Department of Computer Science at the University of British Columbia, where she designed and evaluated a computational tutor that supported students during analogical problem solving. She is a Research Scientist in the Department of Computing, Informatics, and Decision Systems Engineering at Arizona State University. Her work falls into the intersection of Human-Computer-Interaction and Artificial Intelligence, dealing with the design and evaluation of interactive educational technologies that aim to help students learn effectively though personalized support. She is particularly interested in technologies that support high level student states related to meta-cognition, affect, and creativity.

Dr. Tom Murray (USA) is a Senior Research Fellow in School of Computer Science at the University of Massachusetts Amherst. His current research areas include supporting social deliberative skills in online contexts, and text analytics for cognitive developmental levels. He has also published in the areas of ITS authoring tools, adaptive hypermedia, intelligent learning environments, and knowledge engineering. He is also publishes papers in the field of integral theory on embodied epistemology, contemplative dialogue practices, and applied ethics. Murray has degrees in educational technology (EdD, MEd), computer science (MS), and physics (BS). He is on the editorial review boards of two international journals, *the International Journal of Artificial Intelligence in Education* and *Integral Review* (as an Associate Editor).

Dr. Rodney Nielsen (USA) received a dual Ph.D. in computer science and cognitive science from the University of Colorado, Boulder in 2008. He is currently an associate professor of computer science and engineering at UNT, where he co-directs the Language and Information Technologies (LIT) lab. Prior to UNT, he was an assistant professor adjunct of computer science at CU Boulder, a research scientist in CU's Center for Computational Language and Education Research, and a research scientist with Boulder Language Technologies. Dr. Nielsen's research includes the areas of natural language processing, machine learning, and cognitive science, with an emphasis on educational technology (classroom engagement technology and intelligent tutoring systems), spoken-dialogue educational health and wellbeing companion robots (companionbots), health and clinical informatics, and end-user software engineering. One common theme across all of these areas is the need for a robust, informative model of the humans interacting with the system. Further information regarding Dr. Nielsen's research and contact information can be found at http://www.cse.unt.edu/~nielsen/.

Dr. Richard Noss (GBR) is the founder of the London Knowledge Lab and its co-director for its first ten years. He is Professor of Mathematics Education, holding a Masters degree in pure mathematics and a PhD in mathematical education. He was co-founder and deputy scientific manager of Kaleidoscope, the European network of excellence for technology enhanced learning, and was until recently the director of the UK's Technology Enhanced Learning Research Programme funded jointly by the ESRC and EPSRC. His research has focused on the design of constructionist computational environments for learning a range of ideas, mostly mathematics-related. He is a founding member of the International Journal of Computers for Mathematical

Learning. He has also extensively researched the kinds of mathematical knowledge needed by employees in technology-rich workplaces, and appropriate ways to harness technology to foster this knowledge. Richard is a Fellow of the Institute of Mathematics and its Applications, an Academician of the Academy of the Social Sciences and in 2011, was elected a Foreign Fellow of the Union of Bulgarian Mathematicians. He is, since 2012, a Professorial Fellow of the University of Melbourne.

Dr. Benjamin D. Nye (USA) is the Director of Learning Sciences at University of Southern California Institute for Creative Technologies (USC-ICT). His major research interest is to identify best practices in advanced learning technology, particularly for frontiers such as distributed learning technologies (e.g., cloud-based, device-agnostic) and socially situated learning (e.g., face-to-face mobile use). Research interests include modular ITS designs, modeling social learning and memes, cognitive agents, and educational tools for the developing world and low-resource/low-income contexts. He received his PhD in systems engineering from the University of Pennsylvania in 2011. In his recent work as a research professor at UofM, he led work on the Shareable Knowledge Objects (SKO) framework integrating ITS services such as AutoTutor for the ONR ITS Grand Challenge, helped data mine effort a corpus of 250k human-to-human online tutoring dialogs (part of the ADL PAL initiative), collaborated on ONR's PAL3 tutoring architecture for supporting life-long learning, and is an advisor and book editor for the ARL Generalized Intelligent Framework for Tutoring (GIFT) expert workshop panel. His research tries to remove barriers development and adoption of ITSs, so that they can reach larger numbers of learners, which has traditionally been a major roadblock for these highly effective interventions. He also believes that the future of learning science depends on large, sustainable platforms with many users, where efficient sampling techniques can be used to drive new designs for experiments. Finally, he is interested in making the process of science more efficient, such as by advanced metadata and analysis for scholarly publications.

Dr. Brent Olde (USA) is a Lieutenant Commander in the US Navy. He is currently assigned as a Program Officer and Division Deputy at ONR's Human & Bio-Engineered Systems Division. He manages several S&T programs; primarily Live, Virtual, and Constructive (LVC) training; Unmanned Aerial Systems (UAS) Selection, Interface, and Training; and STEM ITSs. He received his undergraduate degree at the University of Missouri – Columbia and his PhD in experimental psychology at UofM. Upon completion he was commissioned as a Lieutenant in the US Navy, completed primary flight training in 2003, and was designated an US Navy AEP. He has completed tours at NAVAIR 1.0, Program Manager (PMA205 – Training Systems); NAVAIR 4.6, Human Systems Research and Engineering Department; Naval Postgraduate School, Assistant Professor; and Naval Aerospace Medicine Institute (NAMI), Fleet Support Division Officer.

Dr. Andrew Olney (USA) presently serves as associate professor in both the Institute for Intelligent Systems (IIS) and Department of Psychology and as director of the IIS at University of Memphis (UofM). He received a BA in linguistics with cognitive science from University College London in 1998, an MS in evolutionary and adaptive systems from the University of Sussex in 2001, and a PhD in computer science from the UofM in 2006. His primary research interests are in natural language interfaces. Specific interests include vector space models, dialogue systems, unsupervised grammar induction, robotics, and ITSs.

Dr. Ana Paiva (PRT). Ana Paiva is Assistant Professor at the Informatics Engineering Department of the Faculty of Engineering at the University of Porto (FEUP) where she works since 1999. She is a researcher at INESC TEC in the Software Engineering area and member of the Software Engineering Group which gathers researchers and post graduate students with common interests in software engineering. She teaches subjects like Software Testing, Formal Methods and Software Engineering, among others. She has a PhD in Electrical and Computer Engineering from FEUP with a thesis titled "Automated Specification Based Testing of Graphical User Interfaces". Her expertise is on the implementation and automation of the model based testing process.

She has been developing research work in collaboration with Foundation of Software Engineering research group within Microsoft Research where she had the opportunity to extend Microsoft's model-based testing tool, Spec Explorer, for GUI testing. She is PI of a National Science Foundation funded project on Pattern-Based GUI Testing (PBGT). She is a member of the PSTQB (Portuguese Software Testing Qualification Board) board general assembly, member of TBok, Glossary, and Examination Working Groups of the ISTQB (International Software Testing Qualification Board), and member of the Council of the Department of Informatics Engineering of FEUP.

Dr. Philip I. Pavlik (USA) is an assistant professor and Director of the Optimal Learning Lab. The lab's mission is to describe models of learning so that these models can be used by instructional software to sequence and schedule practice. He completed his dissertation research with John Anderson in CMU's Psychology Department and has worked with Ken Koedinger in CMU's Human Computer Interaction Institute. He is current working on multiple existing grants and has applied for funding from both Department of Education (DOED) and NSF.

Dr. Radek Pelanek (CZE) is an associate professor in the Department of Information Technologies, Faculty of Informatics at Masaryk University in the Czech Republic. His research interests include adaptive learning, intelligent tutoring systems, educational data mining, learning analytics, and computational modeling. Dr. Pelanek's adaptive learning group has developed a number of adaptive practice systems including an application for learning geography which features an open learner model.

Professor Alex "Sandy" Pentland (USA) directs the MIT Connection Science and Human Dynamics labs and previously helped create and direct the MIT Media Lab and the Media Lab Asia in India. He is one of the most-cited scientists in the world, and Forbes recently declared him one of the "7 most powerful data scientists in the world" along with Google founders and the Chief Technical Officer of the United States. He has received numerous awards and prizes such as the McKinsey Award from Harvard Business Review, the 40th Anniversary of the Internet from DARPA, and the Brandeis Award for work in privacy. He is a founding member of advisory boards for Google, AT&T, Nissan, and the UN Secretary General, a serial entrepreneur who has co-founded more than a dozen companies including social enterprises such as the Data Transparency Lab, the Harvard-ODI-MIT DataPop Alliance and the Institute for Data Driven Design. He is a member of the U.S. National Academy of Engineering and leader within the World Economic Forum. Over the years Sandy has advised more than 60 PhD students. Almost half are now tenured faculty at leading institutions, with another one-quarter leading industry research groups and a final quarter founders of their own companies. Together Sandy and his students have pioneered computational social science, organizational engineering, wearable computing (Google Glass), image understanding, and modern biometrics. His most recent books are `Social Physics,' published by Penguin Press, and 'Honest Signals', published by MIT Press. Interesting experiences include dining with British Royalty and the President of India, staging fashion shows in Paris, Tokyo, and New York, and developing a method for counting beavers from space.

Dr. Ray S. Perez (USA) is a senior scientist and Program Officer at the Office of Naval Research in Arlington, VA. In this capacity, he is responsible for managing ONR's Cognitive Science of Learning Program. This program has three major interdisciplinary and highly intertwined thrusts. Specifically, he is responsible for:

1) Training/education research and their core technologies;
2) Individual differences research; and
3) Neuro-biology of learning research.

He also serves as the Service training lead for the Human Systems Community of Interest for DoD. Dr. Perez's research in the areas of technology-based education and training spans over 30 years. Throughout his career, he has received numerous awards for his work in advanced learning technologies. He has edit six books on the use of technology in education and training. Recently, he co-edited a book with Drs. O'Neil and Baker entitled *Teaching and Measuring Cognitive Readiness* published in 2014 by Springer. Prior to coming to ONR he served as program manager for the Presidential Technology Initiative Program at the Department of Defense Education Activity (DoDEA). While at DoDEA he was the Director of the K–12 program within the Advanced Distribute Learning Initiative, sponsored by the Office of the Secretary of Defense, Readiness and Training. Earlier, he was principal scientist in Simulation and Advanced Instructional Systems, at the U.S. Army Research Institute for the Social and Behavioral Sciences (ARI) and was an assistant professor, in the Department of Psychology, at California State University Dominguez Hills, California. Dr. Perez continues to serve as an educational technology expert on various review panels including the National Science Foundation (NSF), National Academy Sciences (NAS), and the Defense Advanced Research Agency (DARPA). Dr. Perez received a Doctorate and Master's degrees in Educational Psychology with emphasis on Cognitive Psychology from the University of California, Los Angeles California. Dr. Perez is currently the co-chair of NATO Human Factors and Medicine Panel's HFM-237 Research Task Group on the "Assessment of Intelligent Tutoring System Technologies and Opportunities".

Dr. Octav Popescu (USA) is a Senior Research Programmer/Analyst in CMU's Human-Computer Interaction Institute, where he is in charge of TutorShop, the learning management system part of the CTAT project. He has more than 25 years of experience working on various projects involving natural language understanding and ITSs. He holds an MS in computational linguistics and a PhD in language technologies from CMU.

Dr. Anna N. Rafferty (USA) is an assistant professor at Carleton College. Her research focuses on using machine learning and computational cognitive science to build more effective personalized educational technologies. She is particularly interested in how probabilistic models of cognition can be leveraged to provide more effective feedback to learners and to adapt the experience of learners within an educational technology. She has also collaborated with groups focused on science education, such as the Web-based Inquiry Science Environment (WISE) and WestEd, to develop analytics and personalized feedback for chemistry activities used in the classroom. She received her doctorate in computer science from the University of California, Berkeley, and a Master's degree in symbolic systems from Stanford University.

Dr. Sowmya Ramachandran (USA) is a Research Scientist at Stottler Henke Associates where her work focuses on the application of AI and machine learning to improve education and training. She leads research and development of ITSs and ITS authoring tools for a diverse range of military and civilian domains. Dr. Ramachandran headed the development of ReadInsight, an intelligent tutor for teaching reading comprehension skills to adult English speakers. She also led the development of an intelligent tutor for training Tactical Action Officers in the Navy. This system uses NLP technologies to assess and train Tactical Action Officers (TAOs) and is currently in operational use at the Surface Warfare Officers School. She is currently leading the development of an ITS for training US Navy Information Systems Technicians in troubleshooting and maintenance skills. Dr. Ramachandran holds a PhD from The University of Texas at Austin. For her dissertation, she developed a novel machine learning technique for constructing Bayesian network models from data.

Dr. Mark Riedl (USA) is an Associate Professor in the Georgia Tech School of Interactive Computing and director of the Entertainment Intelligence Lab. Dr. Riedl's research focuses on the intersection of artificial intelligence, storytelling, and virtual worlds. Dr. Riedl seeks to understand how computational systems can represent, reason about, and create narratives and interactive stories. His primary research is in automated narrative generation, the creation of fictional narratives by intelligent systems. He also explores how intelligent systems can improve human experiences in games and virtual worlds through dynamic game adaptation and automated game design. Dr. Riedl earned a PhD degree in 2004 from North Carolina State University, where he developed intelligent systems for generating stories and managing interactive user experiences in computer games. From 2004 to 2007, Dr. Riedl was a Research Scientist at the University of Southern California Institute for Creative Technologies where he researched and developed interactive, narrative-based training systems. Dr. Riedl joined the Georgia Tech College of Computing in 2007 and in 2011 he received a DARPA Young Faculty Award and NSF CAREER Award for his work on artificial intelligence, narrative, and virtual worlds.

Dr. Steven Ritter (USA) is the founder and chief scientist at Carnegie Learning where he has been developing and evaluating educational systems for over 20 years. He earned his PhD in cognitive psychology at CMU in 1992 and was instrumental in the development and evaluation of Cognitive Tutors for mathematics. Through leadership of Carnegie Learning's research department, he has led many improvements to the use of adaptive learning systems and math education in real-world settings. He is the author of numerous papers on the design, architecture, and evaluation of ITSs. He is lead author of an evaluation judged by the DOED's What Works Clearinghouse as fully meeting their standards and is lead author of a "Best Paper" at the International Conference on Educational Data Mining.

Dr. Robby Robson (USA) is a learning technology researcher, innovator, and entrepreneur. His first career was in mathematics in the areas of real algebraic geometry and computational number theory. In 1995, he began developing web-based learning environments and technologies, including one of the first online calculus courses. Soon thereafter he became involved in developing industry standards for eLearning interoperability, contributing to several of the early IMS and IEEE standards and chairing the IEEE Learning Technology Standards Committee from 2000 – 2008. Dr. Robson co-founded Eduworks in 2001 where he is CEO and chief scientist. Over the past dozen years he guided Eduworks while helping numerous commercial, academic, and non-profit organizations design and develop educational and training technology and formulate market strategies while serving as the principle investigator or lead scientist on multiple National Science Foundation and Department of Defense projects related to repositories, open content formats, competency management, and intelligent tutoring systems. His most recent work is in applications of semantic technology and natural language processing. He holds a doctorate from Stanford University in mathematics and has held posts in both academia and industry.

Dr. Ido Roll (CAN) is the Senior Manager for Research and Evaluation in the Centre for Teaching, Learning, and Technology at the University of British Columbia (UBC), and he is a researcher with the Pittsburgh Science of Learning Centre. Ido graduated from the Human-Computer Interaction Institute and the Program for Interdisciplinary Education Research in Carnegie Mellon University. Ido studies how interactive learning environments support students in becoming more competent, curious, creative, and collaborative learners in classroom and online environments. His work focuses on cognitive and non-cognitive factors across different time scales, from minutes (in problem solving environments and simulations) to months (in MOOCs and learning management systems). His research utilizes a variety of methodologies from the fields of cognitive science, the learning sciences, artificial intelligence, learning analytics, education, and human-computer interaction. His publications in these fields have won numerous awards, and his research has been funded by the National Science Foundation (NSF), the Social Sciences and Humanities Research Council of Canada (SSHRC), the Gordon and Betty Moore Foundations (GBMF), and others.

Dr. Carolyn Rose (USA) is a Research Scientist with a 50/50 joint appointment between the Language Technologies Institute and the Human-Computer Interaction Institute at Carnegie Mellon University. She earned her PhD in Language and Information Technologies from Carnegie Mellon in 1997. She then worked as a

Research Associate at the Learning Research and Development Center for 6 years working on tutorial dialogue systems. She has been at Carnegie Mellon as a faculty member since Fall of 2003. Carolyn Rose's primary research objective is to develop and apply advanced interactive technology to enable effective computer based and computer supported instruction. A particular focus of her research is in exploring the role of explanation and language communication in learning. Thus, one major thrust of her research is in developing and applying language technology to the problem of eliciting, responding to, and automatically analyzing student verbal behavior. However, many of the underlying HCI issues that are central to her work, such as influencing student expectations, motivation, and learning orientation, transcend the specific input modality.

Dr. Vasile Rus (USA) is an associate professor of computer science at the University of Memphis with a joint appointment in the Institute for Intelligent Systems (IIS) whose areas of expertise are computational linguistics, artificial intelligence, software engineering, and computer science in general. His research areas of interest include question answering and asking, dialogue-based intelligent tutoring systems (ITSs), knowledge representation and reasoning, information retrieval, and machine learning. For the past 10 years, Dr. Rus has been heavily involved in various dialogue-based ITS projects including systems that tutor students on science topics (DeepTutor), reading strategies (iSTART), writing strategies (W-Pal), and metacognitive skills (MetaTutor). Currently, Dr. Rus leads the development of the first intelligent tutoring system based on learning progressions, DeepTutor (www.deeptutor.org). He has coedited two books, received several Best Paper Awards, and authored more than 90 publications in top, peer-reviewed international conferences and journals. He is currently Associate Editor of the *International Journal on Artificial Intelligence Tools*.

Dr. Eduardo Salas (USA) is full professor at Rice University and has co-authored over 489 journal articles and book chapters, and has co-edited over 25 books. He is on/has been on the editorial boards of Personnel Psychology, Theoretical Issues in Ergonomics Science, Applied Psychology: An International Journal, International Journal of Aviation Psychology, Group Dynamics, The Leadership Quarterly, Journal of Occupational and Organizational Psychology, and several others. He is past Editor of Human Factors journal and current Associated Editor for the Journal of Applied Psychology and Military Psychology. His expertise includes helping organizations on how to foster teamwork, design and implement team training strategies, facilitate training effectiveness, manage decision making under stress, develop performance measurement tools, and design learning and simulation-based environments. His research in teamwork and collective training are paving the way to future collective tutoring capabilities.

Dr. Sae Schatz (USA) is the Director of US Office of the Secretary of Defense's Advanced Distributed Learning (ADL) Initiative. Dr. Schatz is widely respected in the distributed learning community for her work in learning science, with a focus on andragogy, human performance assessment, adaptive training, simulation, online/blended learning, and human-systems integration. Dr. Schatz earned her doctorate in Modeling and Simulation at the University of Central Florida. As director of the ADL Initiative, Dr. Schatz plans to emphasize advocacy for learning science, demonstration of the "art of the possible" involving learning technologies, and wide-ranging collaboration with ADL stakeholders. Prior to joining the ADL Initiative, Schatz was a senior consultant with Executive Development Associates, and before that she was the chief scientist for MESH Solutions, LLC (a DSCI Company) in Orlando, Florida. She has worked on a number of well-known projects including the Marine Corps' "Making Good Instructors Great" effort, the NTSA award-winning Border Hunter research and instructional design project, and the Joint Staff's (J7) Blended Learning–Training System. She served on the UCF faculty from 2006-2011, teaching courses in human-systems integration, visual and web design, and web development, and continues to support the Modeling & Simulation graduate program as an occasional adjunct instructor.

Dr. Dylan Schmorrow (USA) is the Chief Scientist at Soar Technology (SoarTech) where he is leading the advancement of research and technology tracks to build intelligent systems for defense, government, and commercial applications that emulate human decision making in order to make people more prepared, more informed and more capable. He has led numerous initiatives that transformed promising technologies into operational capabilities and he successfully transitioned several significant prototypes to operational use. His past service includes the Deputy Director, Human Performance, Training, and BioSystems at the Office of the Secretary of Defense, Program Manager for DARPA, Research Scientist and Branch Head at the Naval Air Warfare Center, Chief Scientist for Human-Technology Integration at the Naval Research Lab, Assistant Professor at the Naval Postgraduate School, Program Officer at ONR, and Executive Assistant to the Chief of Naval Research. He received a commission in the US Navy in 1993 as a Naval aerospace experimental psychologist and completed naval flight training in 1994. He retired as a US Navy Captain in 2013 after twenty years of service where he was both an aerospace experimental psychologist and an acquisition professional leading research and development programs.

Dr. Wolfgang Schnotz (DEU) is a Professor of General and Educational Psychology at University of Koblenz-Landau. His focus in teaching is on cognitive psychology and instructional psychology. He also teaches language and cognition as well as visualization with a focus on new media. Dr. Schnotz received his PhD from the Technical University Berlin. He held positions at University of Tübingen, University of Bielefeld, University of Vienna, and University of Jena. He is now the head of the Department of General and Educational Psychology, the head of the Multimedia Research Centre and the head of the (German-Science-Foundation supported) Graduate School on Teaching and Learning Processes at the University of Koblenz-Landau. Dr. Schnotz was Chief Editor of the international journal *Learning and Instruction*, member of the International Reading Expert Group for PISA 2009 and is editorial board member of numerous journals. He has published widely in the field of reading and listening comprehension, learning from text, comprehension of graphics, learning with hypermedia and learning from animation. He runs currently various research projects on text-picture-integration skills and coherence formation from conflicting information funded by the German Science Foundation.

Dr. Christian Schunn (USA) is a Professor of Psychology, Senior Scientist at the Learning Research and Development Center, Professor of Intelligent Systems Program, Professor, Learning Sciences and Policy, School of Education at the University of Pittsburgh. Christian Schunn began his education at McGill University in Psychology and Computer Science and earned his doctorate from Carnegie Mellon University. His research interests include: STEM reasoning and learning, design-based learning, web-based peer interaction and instruction, neuroscience of complex learning and innovative cognition. Dr. Schunn has been affiliated with the development of the following research products: SWoRD, Feedback.net, and Goodness-of-Fit. He currently serves on the Board of the following Journal publications: Cognitive Science, Educational Designer, International Journal of STEM Education, and the Journal of Educational Psychology.

Dr. David Williamson Shaffer (USA) is a Professor at the University of Wisconsin-Madison in the Department of Educational Psychology and a Game Scientist at the Wisconsin Center for Education Research. Before coming to the University of Wisconsin, Dr. Shaffer taught grades 4-12 in the United States and abroad, including 2 years working with the Asian Development Bank and US Peace Corps in Nepal. His MS and PhD are from the Media Laboratory at MIT, and he taught in the Technology and Education Program at the Harvard Graduate School of Education. Dr. Shaffer was a 2008–2009 European Union Marie Curie Fellow. He studies how new technologies change the way people think and learn, and his most recent book is *How Computer Games Help Children Learn*.

Dr. Valerie Shute (USA) is the Mack & Effie Campbell Tyner Endowed Professor in Education in the Department of Educational Psychology and Learning Systems at Florida State University. Before coming to FSU in 2007, she was a principal research scientist at Educational Testing Service where she was involved with basic and applied research projects related to assessment, cognitive diagnosis, and learning from advanced instructional systems. Her general research interests hover around the design, development, and evaluation of advanced systems to support learning – particularly related to 21st century competencies. An example of current research involves using immersive games with stealth assessment to support learning – of cognitive and non-cognitive knowledge, skills, and dispositions. Her research has resulted in numerous grants, journal articles, books, chapters in edited books, a patent, and a 2010 book co-edited with Betsy Becker entitled, *Innovative assessment for the 21st century: Supporting educational needs.*

Dr. Anne M. Sinatra (USA) is an adaptive training scientist at the US Army Research Laboratory in Orlando, Florida. The focus of her research is in cognitive and human factors psychology. She has specific interest in how information relating to the self and about those that one is familiar with can aid in memory, recall, and tutoring. Her dissertation research evaluated the impact of using degraded speech and a familiar story on attention/recall in a dichotic listening task. Her work has been published in the Journal of Interaction Studies, and in the conference proceedings of the Human Factors and Ergonomics Society. Prior to becoming an ARL Scientist, Dr. Sinatra was an ARL Post-Doctoral Fellow and Graduate Research Associate with UCF's Applied Cognition and Technology (ACAT) Lab, and taught a variety of undergraduate Psychology courses. Dr. Sinatra received her Ph.D. and M.A. in Applied Experimental and Human Factors Psychology, as well as her B.S. in Psychology from the University of Central Florida.

Dr. Robert E. Slavin (USA) is director of the Center for Research and Reform in Education at Johns Hopkins University, a part-time professor at the Institute for Effective Education at the University of York (England), and chairman of the Success for All Foundation. He received his B.A. in psychology from Reed College in 1972 and his Ph.D. in social relations in 1975 from Johns Hopkins University. He has authored or co-authored more than 300 articles and book chapters on such topics as cooperative learning, comprehensive school reform, ability grouping, school and classroom organization, desegregation, mainstreaming, research review and evidence-based reform. He is the author or co-author of 24 books, including Educational Psychology: Theory into Practice (Allyn & Bacon, 1986, 1988, 1991, 1994, 1997, 2000, 2003, 2006, 2009); Cooperative Learning: Theory, Research, and Practice (Allyn & Bacon, 1990, 1995); Show Me the Evidence: Proven and Promising Programs for America's Schools (Corwin, 1998); Effective Programs for Latino Students (Erlbaum, 2000);Educational Research in the Age of Accountability (Allyn & Bacon, 2007); and Two Million Children: Success for All (Corwin, 2009). He received the American Educational Research Association's Raymond B. Cattell Early Career Award for Programmatic Research in 1986; Palmer O. Johnson award for the best article in an AERA journal in 1988; Charles A. Dana award in 1994; James Bryant Conant Award from the Education Commission of the States in 1998; Outstanding Leadership in Education Award from the Horace Mann League in 1999; Distinguished Services Award from the Council of Chief State School Officers in 2000; AERA Review of Research Award in 2009; and Palmer O. Johnson Award for the best article in an AERA journal in 2008; and was appointed as a member of the National Academy of Education in 2009 and an AERA fellow in 2010.

Dr. Robert A. Sottilare (USA) is a Senior Scientist at the US Army Research Laboratory with over 30 years of experience in the modeling, simulation, and training domain. The focus of his research is automated authoring, instructional management, and analysis tools and methods for intelligent tutoring systems (ITSs). As an ARL scientist, he leads the adaptive training research program and its affiliated Learning in Intelligent Tutoring Environments (LITE) laboratory, and is the founder of the Center for Adaptive Instructional Sciences (CAIS). He is a co-creator of the Generalized Intelligent Framework for Tutoring (GIFT) and lead editor for the *Design Recommendations for Intelligent Tutoring* book series (currently four volumes). He is a member of the *AI in*

Education Society. His work is widely published (over 130 technical publications) and includes articles in the *Journal for Defense Modeling & Simulation, Cognitive Technology and the Educational Technology Journal & Society*. He has served in several leadership positions in support of technical conferences including recent roles at: Augmented Cognition (2014, 2015), Defense & Homeland Security Simulation (2011 – 2012, 2015 –2016), and the Florida Artificial Intelligence Research Society (2014 – 2016). Dr. Sottilare also contributes time as a technical reviewer for several journals including the *Theoretical Issues in Ergonomic Sciences (TIES) Journal* and the *International Journal of Human Factors Modeling & Simulation*. He is a faculty scholar and lecturer at the United States Military Academy (USMA) and adjunct professor at the University of Central Florida (UCF) in the Modeling & Simulation Program. Dr. Sottilare is a recipient of two lifetime achievement awards in Modeling & Simulation: *US Army RDECOM* (2012) and *National Training & Simulation Association* (2015). Dr. Sottilare is active in NATO science and technology activities and as of this date is the co-chair of NATO Human Factors and Medicine Panel's HFM-237 Research Task Group on the "Assessment of Intelligent Tutoring System Technologies and Opportunities".

Dr. Gerry Stahl (USA) is an Emeritus Professor of the College of Computing and Informatics at Drexel University in Philadelphia, PA. He is a researcher and professor emeritus of information science. His current research (from 2003 to the present) focuses on the Virtual Math Teams (VMT) project at Drexel University's College of Computing and Informatics, the Math Forum and Rutgers-Newark. This project is extensively documented in my book on *Studying Virtual Math Teams*. Our research team uses chat interaction analysis to explore what takes place in online discussion of math by students. The background and motivation for this research in my previous research from 1990 to 2005 is presented in my book on Group Cognition. His recent (2013) book, Translating Euclid, discusses the redesign of geometry education in terms of: cognitive history, contemporary philosophy, school mathematics, software technology, collaborative learning, design-based research, CSCL theory, developmental pedagogy and scaffolded practice. It is a multi-dimensional reflection on the VMT Project, including the very latest ideas and findings. His new book (2015), Constructing Dynamic Triangles Together, analyzes how a team of three junior-high-school girls developed their mathematical group cognition through eight hour-long online sessions using VMT to learn dynamic geometry. His specialty is Computer-Supported Collaborative Learning (CSCL).

Dr. Slavomir Stankov (HRV) is a full professor in the Department of Computer Science at the University of Split in Croatia. His research interests include ITSs, e-learning, e-learning systems, authoring systems, software engineering. Dr. Stankov has several projects related to ITS science and design, but is most noted for his creation of *TexSys*, Tutor–Expert System model for building ITSs systems in freely chosen task domains.

Dr. Ronald H. Stevens (USA) is Professor (Emeritus), UCLA School of Medicine and a member of the UCLA Brain Research Institute. His recent research has focused on using EEG-derived measures to investigate Team NeuroDynamics in settings as diverse as Submarine Piloting and Navigation by Navy teams and high school problem solving. These studies are leading to quantitative teamwork models showing how teams cognitively organize in response to environmental and task changes. Recent awards include the Foundations of Augmented Cognition Award (2007), Best Paper, Behavioral Research in Modeling and Simulation Conference (2011), and the Admiral Leland Kollmorgen Spirit of Innovation Award from the Augmented Cognition Group (2012). The research support has been supported by the National Science Foundation, DARPA, Office of Naval Research, Dept. of Education and multiple private donors.

Dr. Choo-Yee Ting (MYS) is currently Associate Professor in the Faculty of Computing and Informatics at Multimedia University in Cyberjaya, Malaysia. In 2002, Dr. Ting was awarded the Fellow of Microsoft Research by Microsoft Research Asia, Beijing, China. In 2003, he received a research fellowship from the Rotary Research Foundation, Rotary Club of Kuala Lumpur Diraja, Malaysia. He has been involved in research projects funded by MOSTI, Malaysia and Industries. He is also certified in Microsoft Technology Associate (Database) and IBM DB2 CDA. His work in multi-agent AI techniques in Intelligent Learning Systems, and INQPRO, an intelligent scientific inquiry exploratory learning environment are noteworthy.

Dr. David Traum (USA) is a principal scientist at the Institute for Creative Technologies (ICT), and research faculty in the Computer Science Department, both at the University of Southern California. At ICT, Traum leads the Natural Language Dialogue Group (http://projects.ict.usc.edu/nld/group/). Traum's research focuses on dialogue communication between human and artificial agents. He has engaged in theoretical, implementational and empirical approaches to the problem, studying human-human natural language and multi-modal dialogue, as well as building a number of dialogue systems to communicate with human users. He has pioneered several research thrusts in computational dialogue modeling, including computational models of grounding (how common ground is established through conversation), the information state approach to dialogue, multiparty dialogue, and non-cooperative dialogue. Dr. Traum is author of over 200 technical articles, is a founding editor of the *Journal Dialogue and Discourse* has chaired and served on many conference program committees, and is on the board and a past president of SIGDIAL, the international special interest group in discourse and dialogue. He earned his Ph.D. in computer science at University of Rochester in 1994.

Dr. Kurt VanLehn (USA) is the Diane and Gary Tooker Chair for Effective Education in Science, Technology, Engineering and Math in the Ira A. Fulton Schools of Engineering at Arizona State University. He received a Ph.D. from MIT in 1983 in computer science, was a post-doc at BBN and Xerox PARC, joined the faculty of CMU in 1985, moved to the University of Pittsburgh in 1990 and joined ASU in 2008. He founded and co-directed two large NSF research centers (Circle; the Pittsburgh Science of Learning Center). He has published over 125 peer-reviewed publications, is a fellow in the Cognitive Science Society, and is on the editorial boards of *Cognition and Instruction* and the *International Journal of Artificial Intelligence in Education*. Dr. VanLehn's research focuses on intelligent tutoring systems and other intelligent interactive instructional technology.

Dr. William Vessey (USA) is a portfolio scientist for the Team Risk within the NASA Behavioral Health and Performance Research Element at the Johnson Space Center. He received an MS and PhD in Industrial and Organizational Psychology from the University of Oklahoma in 2012 with a minor in Quantitative Psychology. His primary research interests fall into the broad categories of teams, leadership, and creativity with specific focus on teamwork over long durations, team leadership and collective leadership. His work has appeared in several books and journals, including The Leadership Quarterly, Creativity Research Journal, Creativity and Innovation Management, The Encyclopedia of Creativity, and Leadership 101.

Dr. Wayne Ward (USA) is Principal Scientist and Chief Financial Officer at Boulder Language Technologies, Inc. He received a BA in mathematical science and psychology at Rice University and an MA and PhD in psychology at University of Colorado. He is also a Research Professor at the Computational Language and Education Research Center for the University of Colorado. Previously, he was appointed as a Research Computer Scientist at CMU. Dr. Ward developed and maintains the Phoenix system, a parser and dialogue manager designed specifically for semantic information extraction from spoken dialogues in limited domains. Phoenix is distributed as freeware by Boulder Language Technologies and by CMU. Dr. Ward then led the effort to incorporate these and additional technologies into the Virtual Human Toolkit, a toolkit for developing conversational systems using animated agents.

Dr. Joseph Jay Williams (USA) is a research fellow at Harvard University's Office of the Vice Provost for Advances in Learning. He is also affiliated with the Intelligent Interactive Systems Group in Harvard Computer Science, and the ASSISTments K12 educational platform at Worcester Polytechnic Institute. His research designs adaptive systems for online content, by integrating research in psychology and education, human-computer interaction, and statistical machine learning. To make any static website become intelligently adaptive, he uses powerful systems for randomized A/B experiments. Examples range from adaptive explanations for how to solve math problems, to self-personalizing emails that change people's behavior. These systems continually crowdsource new "A" and "B" designs from psychological scientists, using randomized comparisons to evaluate how helpful these alternative designs are, for people with different profiles. Algorithms from statistical machine learning use this data for real-time enhancement and personalization, by changing which designs are presented to future users. He completed a postdoc in 2014 at Stanford University's Graduate School of Education. He received his PhD in 2013 in from UC Berkeley's Psychology Department, where his research investigated why explaining "why?" helps learning, and used Bayesian statistics and machine learning to model human cognition. He received a BS from University of Toronto in cognitive science, AI, and mathematics, and is originally from Trinidad and Tobago.

Dr. Eliot Winer (USA) is an associate professor in Mechanical Engineering with courtesy appointment as an associate professor in Electrical and Computer Engineering at Iowa State University. He is also the associate director for the Virtual Reality Applications Center at Iowa State. Dr. Winer received his PhD at the University of Buffalo in Mechanical Engineering in 1999. His research interests include Internet Technology for Large-Scale Collaborative Design, Medical Imaging Analysis and Visualization, Multidisciplinary Design Synthesis, Computer Aided Design and Graphics, Applications in Optimal Design, Scientific Visualization, and Virtual Reality Modeling for Large-Scale Design. He is currently involved in an adaptive team tutoring project for the US Army Research Laboratory.

Dr. Beverly Park Woolf (USA) is a research professor at the University of Massachusetts who develops intelligent tutors that model student affective and cognitive characteristics and combine cognitive analysis of learning with artificial intelligence, network technology, and multimedia. These systems represent the knowledge taught, recognize learners' skills and behavior, use sensors, and machine learning to model student affect, and adjust problems to help individual students. Dr. Woolf has developed tutors in education and industry and in a variety of disciplines (e.g., chemistry, psychology, physics, geology, art history, mathematics, and economics). Some of these tutors enable students to pass standard exams at a 20% higher rate and one system is used by more than 150,000 students per semester across hundreds of colleges. Dr. Woolf published the book Building Intelligent Interactive Tutors along with over 200 articles. She is lead author on the NSF report Roadmap to Education Technology in which 40 experts and visionaries identified the next big computing ideas that will define education technology and developed a vision of how technology can incorporate deeper knowledge about human cognition and develop dramatically more effective instructional strategies. Dr. Woolf has delivered keynote addresses, panels, and tutorials in more than 20 foreign countries and is a fellow of the American Association of Artificial Intelligence.

Dr. R. Michael Young (USA) is a professor of computer science at North Carolina State University, where he leads the Liquid Narrative Research Group. He's the founder and director of the NCSU Digital Games Research Center. His work focuses on the computational modeling of interactive narrative, especially in the context of computer games and virtual worlds. He teaches courses on game development and interactive narrative. In 2000, Michael received a CAREER Award from the US National Science Foundation. He has received awards for both outstanding teaching and outstanding activities in economic development. In 2010, Michael was awarded a GlaxoSmithKline Faculty Fellowship for Public Policy and Public Engagement. Michael is a senior member of IEEE and of AAAI and an ACM Distinguished Scientist. Michael was co-founder of several conferences that are leading venues for publication of scientific advances in computer games. Michael was editor-in-chief of the *Journal of Game Development* from 2007 to 2008 and an associate editor of the ACM journal *Transactions on Interactive Intelligence Systems* from 2012 through 2013. He serves as an associate editor of the IEEE journal *Transaction on Computational Intelligence and AI in Games* and the journal *Advances in Cognitive Systems*.

Dr. Diego Zapata-Rivera (USA) is a Senior Research Scientist in the Cognitive and Learning Sciences Center at ETS in Princeton, NJ. He earned a PhD in computer science (with a focus on AI in education) from the University of Saskatchewan in 2003. His research at ETS has focused on the areas of innovations in score reporting and Technology-Enhanced Assessment (TEA) including work on adaptive learning environments and game-based assessments. His research interests also include evidence-centered design, Bayesian student modeling, open student models, conversation-based tasks, virtual communities, authoring tools, and program evaluation. Dr. Zapata-Rivera has produced over 100 publications including journal articles, book chapters, and technical papers. He has served as a reviewer for several international conferences and journals. He has been a committee member and organizer of international conferences and workshops in his research areas. He is a member of the Board of Special Reviewers of the User Modeling and User-Adapted Interaction journal and an Associate Editor of the *IEEE Transactions on Learning Technologies Journal*. Most recently, Dr. Zapata-Rivera has been invited to contribute his expertise to projects sponsored by the National Research Council, NSF, and NASA.

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's References	3. Further Reference	4. Security Classification of Document
	STO-TR-HFM-237 AC/323(HFM-237)TP/764	ISBN 978-92-837-2091-1	PUBLIC RELEASE

5. Originator Science and Technology Organization
North Atlantic Treaty Organization
BP 25, F-92201 Neuilly-sur-Seine Cedex, France

6. Title

Assessment of Intelligent Tutoring Systems Technologies and Opportunities

7. Presented at/Sponsored by

This Report documents the findings of the Human Factors and Medicine Task Group 237.

8. Author(s)/Editor(s)	9. Date
Multiple	January 2018

10. Author's/Editor's Address	11. Pages
Multiple	140

12. Distribution Statement There are no restrictions on the distribution of this document. Information about the availability of this and other STO unclassified publications is given on the back cover.

13. Keywords/Descriptors

Accelerated learning
Adaptive instruction
Adaptive interfaces
Authoring tools
Data analytics
Data mining
Instructional strategies
Intelligent Tutoring Systems (ITS)
Learner competency
Task domains

14. Abstract

This report summarizes the work and findings of HFM Research Task Group (RTG) HFM-237 which was established to explore the potential of Intelligent Tutoring System (ITS) technologies and identify opportunities for their use in NATO training. This report provides a ready reference of ITS research and technology and includes chapters on ITS authoring, emerging technologies and application domains (collective tutoring; Science, Technology, Engineering and Mathematics – STEM tutoring; and medical tutoring). The RTG examined the literature and activities in countries within and outside of NATO to discover the background, opportunities, and limits of ITS technologies (tools and methods). Expert guest lecturers and a review of the literature, current and emerging research, and prototype development by the RTG identified challenges in four major areas: authoring (development), standardization, data analytics, and adaptive interfaces. Recommendations for adoption of ITS technologies and additional research are provided in each of the four major challenge areas in Chapter 8 of this report.

NORTH ATLANTIC TREATY ORGANIZATION

SCIENCE AND TECHNOLOGY ORGANIZATION

BP 25
F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE
Télécopie 0(1)55.61.22.99 • E-mail mailbox@cso.nato.int

DIFFUSION DES PUBLICATIONS
STO NON CLASSIFIEES

Les publications de l'AGARD, de la RTO et de la STO peuvent parfois être obtenues auprès des centres nationaux de distribution indiqués ci-dessous. Si vous souhaitez recevoir toutes les publications de la STO, ou simplement celles qui concernent certains Panels, vous pouvez demander d'être inclus soit à titre personnel, soit au nom de votre organisation, sur la liste d'envoi.

Les publications de la STO, de la RTO et de l'AGARD sont également en vente auprès des agences de vente indiquées ci-dessous.

Les demandes de documents STO, RTO ou AGARD doivent comporter la dénomination « STO », « RTO » ou « AGARD » selon le cas, suivi du numéro de série. Des informations analogues, telles que le titre est la date de publication sont souhaitables.

Si vous souhaitez recevoir une notification électronique de la disponibilité des rapports de la STO au fur et à mesure de leur publication, vous pouvez consulter notre site Web (http://www.sto.nato.int/) et vous abonner à ce service.

CENTRES DE DIFFUSION NATIONAUX

ALLEMAGNE
Streitkräfteamt / Abteilung III
Fachinformationszentrum der Bundeswehr (FIZBw)
Gorch-Fock-Straße 7, D-53229 Bonn

BELGIQUE
Royal High Institute for Defence – KHID/IRSD/RHID
Management of Scientific & Technological Research for Defence, National STO Coordinator
Royal Military Academy – Campus Renaissance
Renaissancelaan 30, 1000 Bruxelles

BULGARIE
Ministry of Defence
Defence Institute "Prof. Tsvetan Lazarov"
"Tsvetan Lazarov" bul no.2
1592 Sofia

CANADA
DGSIST 2
Recherche et développement pour la défense Canada
60 Moodie Drive (7N-1-F20)
Ottawa, Ontario K1A 0K2

DANEMARK
Danish Acquisition and Logistics Organization (DALO)
Lautrupbjerg 1-5
2750 Ballerup

ESPAGNE
Área de Cooperación Internacional en I+D
SDGPLATIN (DGAM)
C/ Arturo Soria 289
28033 Madrid

ESTONIE
Estonian National Defence College
Centre for Applied Research
Riia str 12
Tartu 51013

ETATS-UNIS
Defense Technical Information Center
8725 John J. Kingman Road
Fort Belvoir, VA 22060-6218

FRANCE
O.N.E.R.A. (ISP)
29, Avenue de la Division Leclerc
BP 72
92322 Châtillon Cedex

GRECE (Correspondant)
Defence Industry & Research General Directorate, Research Directorate
Fakinos Base Camp, S.T.G. 1020
Holargos, Athens

HONGRIE
Hungarian Ministry of Defence
Development and Logistics Agency
P.O.B. 25
H-1885 Budapest

ITALIE
Ten Col Renato NARO
Capo servizio Gestione della Conoscenza
F. Baracca Military Airport "Comparto A"
Via di Centocelle, 301
00175, Rome

LUXEMBOURG
Voir Belgique

NORVEGE
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25
NO-2007 Kjeller

PAYS-BAS
Royal Netherlands Military Academy Library
P.O. Box 90.002
4800 PA Breda

POLOGNE
Centralna Biblioteka Wojskowa
ul. Ostrobramska 109
04-041 Warszawa

PORTUGAL
Estado Maior da Força Aérea
SDFA – Centro de Documentação
Alfragide
P-2720 Amadora

REPUBLIQUE TCHEQUE
Vojenský technický ústav s.p.
CZ Distribution Information Centre
Mladoboleslavská 944
PO Box 18
197 06 Praha 9

ROUMANIE
Romanian National Distribution Centre
Armaments Department
9-11, Drumul Taberei Street
Sector 6
061353 Bucharest

ROYAUME-UNI
Dstl Records Centre
Rm G02, ISAT F, Building 5
Dstl Porton Down
Salisbury SP4 0JQ

SLOVAQUIE
Akadémia ozbrojených síl gen. M.R. Štefánika, Distribučné a informačné stredisko STO
Demänová 393
031 06 Liptovský Mikuláš 6

SLOVENIE
Ministry of Defence
Central Registry for EU & NATO
Vojkova 55
1000 Ljubljana

TURQUIE
Milli Savunma Bakanlığı (MSB)
ARGE ve Teknoloji Dairesi Başkanlığı
06650 Bakanliklar – Ankara

AGENCES DE VENTE

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
ROYAUME-UNI

Canada Institute for Scientific and Technical Information (CISTI)
National Research Council Acquisitions
Montreal Road, Building M-55
Ottawa, Ontario K1A 0S2
CANADA

Les demandes de documents STO, RTO ou AGARD doivent comporter la dénomination « STO », « RTO » ou « AGARD » selon le cas, suivie du numéro de série (par exemple AGARD-AG-315). Des informations analogues, telles que le titre et la date de publication sont souhaitables. Des références bibliographiques complètes ainsi que des résumés des publications STO, RTO et AGARD figurent dans le « NTIS Publications Database » (http://www.ntis.gov).

NORTH ATLANTIC TREATY ORGANIZATION

SCIENCE AND TECHNOLOGY ORGANIZATION

BP 25
F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE
Télécopie 0(1)55.61.22.99 • E-mail mailbox@cso.nato.int

DISTRIBUTION OF UNCLASSIFIED STO PUBLICATIONS

AGARD, RTO & STO publications are sometimes available from the National Distribution Centres listed below. If you wish to receive all STO reports, or just those relating to one or more specific STO Panels, they may be willing to include you (or your Organisation) in their distribution.

STO, RTO and AGARD reports may also be purchased from the Sales Agencies listed below.

Requests for STO, RTO or AGARD documents should include the word 'STO', 'RTO' or 'AGARD', as appropriate, followed by the serial number. Collateral information such as title and publication date is desirable.

If you wish to receive electronic notification of STO reports as they are published, please visit our website (http://www.sto.nato.int/) from where you can register for this service.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Royal High Institute for Defence – KHID/IRSD/RHID
Management of Scientific & Technological Research for Defence, National STO Coordinator
Royal Military Academy – Campus Renaissance
Renaissancelaan 30
1000 Brussels

BULGARIA
Ministry of Defence
Defence Institute "Prof. Tsvetan Lazarov"
"Tsvetan Lazarov" bul no.2
1592 Sofia

CANADA
DSTKIM 2
Defence Research and Development Canada
60 Moodie Drive (7N-1-F20)
Ottawa, Ontario K1A 0K2

CZECH REPUBLIC
Vojenský technický ústav s.p.
CZ Distribution Information Centre
Mladoboleslavská 944
PO Box 18
197 06 Praha 9

DENMARK
Danish Acquisition and Logistics Organization (DALO)
Lautrupbjerg 1-5
2750 Ballerup

ESTONIA
Estonian National Defence College
Centre for Applied Research
Riia str 12
Tartu 51013

FRANCE
O.N.E.R.A. (ISP)
29, Avenue de la Division Leclerc – BP 72
92322 Châtillon Cedex

GERMANY
Streitkräfteamt / Abteilung III
Fachinformationszentrum der Bundeswehr (FIZBw)
Gorch-Fock-Straße 7
D-53229 Bonn

GREECE (Point of Contact)
Defence Industry & Research General Directorate, Research Directorate
Fakinos Base Camp, S.T.G. 1020
Holargos, Athens

HUNGARY
Hungarian Ministry of Defence
Development and Logistics Agency
P.O.B. 25
H-1885 Budapest

ITALY
Ten Col Renato NARO
Capo servizio Gestione della Conoscenza
F. Baracca Military Airport "Comparto A"
Via di Centocelle, 301
00175, Rome

LUXEMBOURG
See Belgium

NETHERLANDS
Royal Netherlands Military Academy Library
P.O. Box 90.002
4800 PA Breda

NORWAY
Norwegian Defence Research Establishment, Attn: Biblioteket
P.O. Box 25
NO-2007 Kjeller

POLAND
Centralna Biblioteka Wojskowa
ul. Ostrobramska 109
04-041 Warszawa

PORTUGAL
Estado Maior da Força Aérea
SDFA – Centro de Documentação
Alfragide
P-2720 Amadora

ROMANIA
Romanian National Distribution Centre
Armaments Department
9-11, Drumul Taberei Street
Sector 6
061353 Bucharest

SLOVAKIA
Akadémia ozbrojených síl gen M.R. Štefánika, Distribučné a informačné stredisko STO
Demänová 393
031 06 Liptovský Mikuláš 6

SLOVENIA
Ministry of Defence
Central Registry for EU & NATO
Vojkova 55
1000 Ljubljana

SPAIN
Área de Cooperación Internacional en I+D
SDGPLATIN (DGAM)
C/ Arturo Soria 289
28033 Madrid

TURKEY
Milli Savunma Bakanlığı (MSB)
ARGE ve Teknoloji Dairesi Başkanlığı
06650 Bakanliklar – Ankara

UNITED KINGDOM
Dstl Records Centre
Rm G02, ISAT F, Building 5
Dstl Porton Down, Salisbury SP4 0JQ

UNITED STATES
Defense Technical Information Center
8725 John J. Kingman Road
Fort Belvoir, VA 22060-6218

SALES AGENCIES

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
UNITED KINGDOM

Canada Institute for Scientific and Technical Information (CISTI)
National Research Council Acquisitions
Montreal Road, Building M-55
Ottawa, Ontario K1A 0S2
CANADA

Requests for STO, RTO or AGARD documents should include the word 'STO', 'RTO' or 'AGARD', as appropriate, followed by the serial number (for example AGARD-AG-315). Collateral information such as title and publication date is desirable. Full bibliographical references and abstracts of STO, RTO and AGARD publications are given in "NTIS Publications Database" (http://www.ntis.gov).

ISBN 978-92-837-2091-1