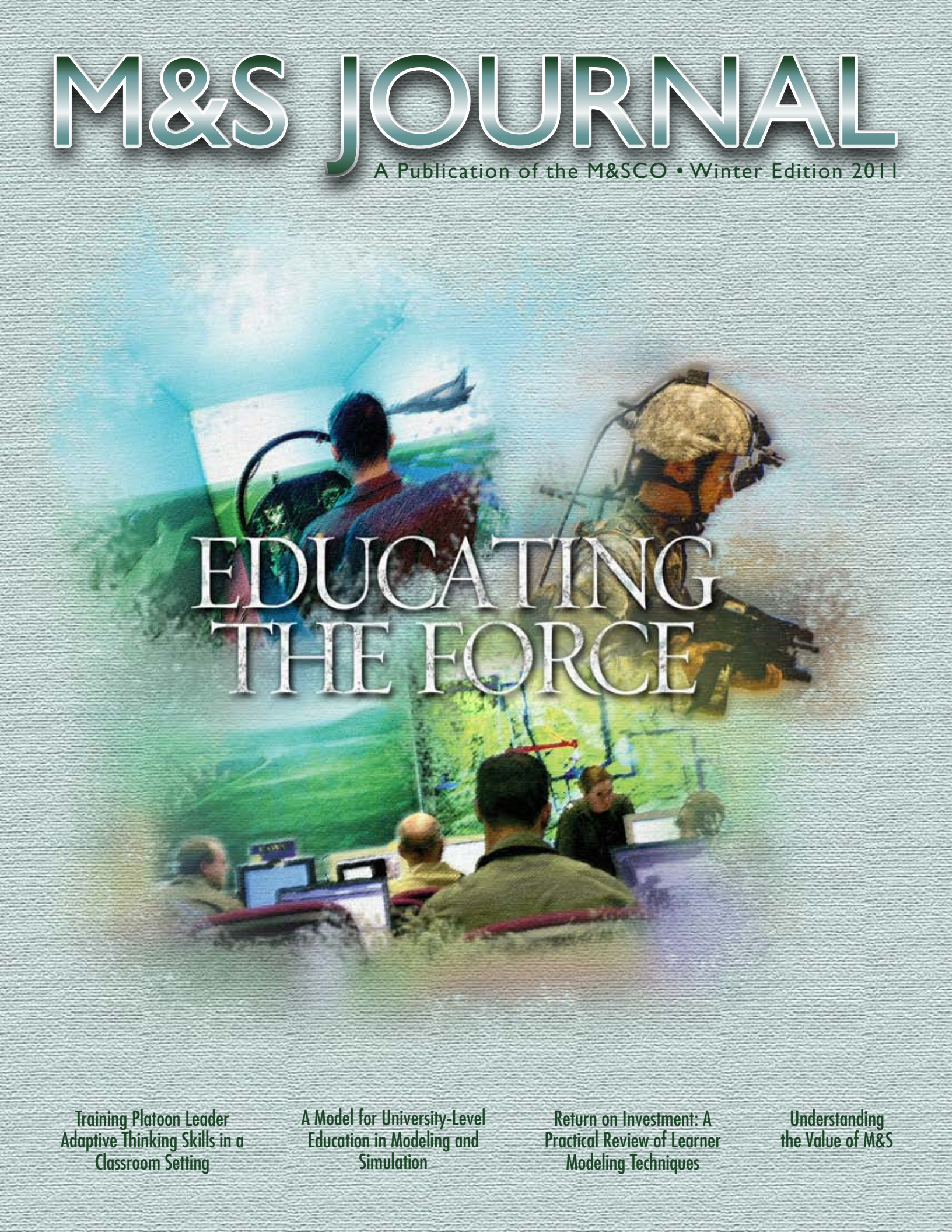
M&S JOURNAL

A Publication of the M&SCO • Winter Edition 2011



Training Platoon Leader Adaptive Thinking Skills in a Classroom Setting

A Model for University-Level Education in Modeling and Simulation

Return on Investment: A Practical Review of Learner Modeling Techniques

Understanding the Value of M&S

FROM THE EXECUTIVE EDITOR

The Department of Defense (DoD) has developed five major goals for Modeling and Simulation (M&S). One of these is educating the workforce. Specifically, this goal calls for people that are well trained; that employ models, simulation, and data to support departmental objectives; and that advance M&S to support emerging departmental challenges. This issue of the M&S Journal is focused on M&S education and its effect on workforce development.

This issue of the M&S Journal contains four informative papers on recent initiatives and strategies relating to educating the force. The article on the "Training Platoon Leader Adaptive Thinking Skills in a Classroom Setting" showcases how M&S education is developing new modalities for advancing workforce development. The next two papers summarize a model for university-level M&S education and review the return on investment for learner modeling techniques. Finally, the last article presents a framework to educate decision makers about the value and limitations of M&S.

Our Guest Editor, Mr. Roger Samuels, Chief of the Army Simulation Proponent and School, provides an editorial on Army M&S Training and Education Programs. Having completed one of these programs, the civilian Career Program 36 (CP36), I can say that specialized knowledge about Simulation Operations, and the integration of M&S and Mission Command systems, has played a critical role in my development as an M&S Professional. For my Developmental Assignment at the Korea Battle Simulation Center (KBSC), US Army Garrison-Yongsan, Seoul, Republic of Korea, I served as Assistant Operations Officer for exercise Key Resolve 2009. There I learned how data, tools, and web services are used to create a realistic operational environment for over 25,000 participants of Joint and Combined theater-level training.

Developing and maintaining a pool of skilled scientists and engineers is critical to achieving our national objectives. The Modeling and Simulation Coordination Office (M&SCO) has established a way forward to guide its investments toward a well trained workforce: encourage the development, coordination and maintenance of M&S training and education programs throughout the DoD; promote M&S education within existing Service and career training programs; establish and maintain partnerships with academia and industry providing M&S education, training, and certification; and promote forums to exchange information and ideas on M&S topics. In addition, M&SCO has an on-going effort to improve the M&S workforce e-learning offerings through the Defense Acquisition University (DAU).

I learned a lot from reading these articles on education in this issue of the M&S Journal. I hope that you do too.

J. DAVID LASHLEE, PH.D., CMSP
Associate Director for Data
Modeling and Simulation Coordination Office
(M&SCO)

Table of Contents

Volume 6 Issue 3

2
Guest Editorial
By:
Roger S. Samuels

5
Training Platoon Leader Adaptive Thinking Skills in a Classroom Setting
By:
Jason Sidman Robert J. Pleban Alexandra Geyer Rob Semmens

14
A Model for University-Level Education in Modeling and Simulation
By:
Roland R. Mielke James F. Leathrum, Jr. Frederic D. McKenzie

24
Return on Investment: A Practical Review of Learner Modeling Techniques
By:
Jeremiah T. Folsom-Kovarik Sae Schatz

38
Understanding the Value of M&S
By:
Jerry M. Feinberg Robert J. Graebener Kriya S. Kaping Michael J. Meehan

46
About the M&S Journal

47
Invitation to Submit an Article

49
Editorial Board

Guest Editorial

Roger S. Samuels
Chief, Simulation Proponent & School
Headquarters, Department of the Army
Deputy Chief of Staff G8, Center for Army Analysis

EDITORIAL

The use of modeling and simulation (M&S) is continually evolving and enhances everything from military readiness, to business processes, to every area of medical care. M&S can provide a pivotal bridge to testing, concepts, and innovations in those instances where either systems do not yet exist, or real life systems cannot be utilized in experimentation or observation because of resource, security, or safety limitations. The effective use of M&S reduces the total life cycle costs of programs across the Services and Industry and the DoD M&S Strategic Vision has reinforced both the need for M&S and a trained and ready workforce. And in July 2007, the U.S. House of Representatives passed Resolution 487, declaring M&S to be a "national critical technology." [1]

In the 1990's, the Army dedicated an officer functional area to the use of M&S. These military professionals, Functional Area 57 Simulation Operations Officers (FA57), have evolved to a force of over 500 positions in the Army's Active, National Guard and Army Reserve components. Throughout the following decade, the Army continued to invest in key M&S resources, and in 2005, established both Civilian Career Program 36 (CP36) and the Army Modeling and Simulation School. Today, the military and civilian programs include close to 3,000 positions across the Army. All three programs are located in the National Capital Area within the Army Simulation Proponent and School, Deputy Chief of Staff, G8, Center for Army Analysis.

Currently within the Army, we are studying the broad range of M&S knowledge, skills and abilities (KSAs) throughout the military and civilian population to better understand the requirements of a full range of jobs associated with M&S. This will include those with very limited KSAs to those at expert levels. This study, when completed, is intended to be exportable and integrated with DoD, Industry and Academia. It will help guide overall resource expenditures in M&S Workforce development. The goal of the study is to:

- Identify and document the skill sets, and types of positions involved with M&S, what they need to know about M&S and what is needed to educate, train and develop them.
- Identify and catalog education, current training and development opportunities for government employees.
- Cross reference knowledge, skills and abilities to specific job types and associated training and education.
- Provide a workforce guide to ensure the workforce can identify M&S training and education opportunities.
- Develop M&S use cases to better facilitate the use of M&S throughout the communities.
- Identify training and education gaps to better synchronize development activities.
- Provide schools, private vendors, certifying organizations, and universities with a more complete picture of the workforce requirements.
- Provide systematic and integrated workforce development through a comprehensive and sustainable development strategy.

M&S also plays an important role in industry, from systems development to product improvements. Its widespread utility, both within the government and in industry, has driven an increased requirement for education, training and development. Within the Army, FA57, CP36 and the Modeling and Simulation School are thriving examples today of the provisions for the vibrant use of M&S and

technology. To further understand the potential of M&S in the DoD requires a closer look into programs like these, and, as understanding increases, it will be clearly evident that further DoD workforce investments are critically necessary.

FUNCTIONAL AREA 57 (FA57)

Simulation Operations FA57 [2] officers are integral to mission success. They provide the interface between the soldier and the technology, providing the Army with a distinct advantage.

These highly trained experts develop simulated environments that support the mission with the development of training exercises based on the Commander's training objectives and requirements. Armed with their operational experience and specialized training, they serve as a critical combat multiplier for the Warfighter. The FA57 officers' further understanding of people, process and technology coupled with their operational experience, provides the Army with a unique capability to manage digital tactical operations centers, disseminate key/critical information and facilitate knowledge transfer. FA57's are invaluable to the Army as M&S experts, mission command integrators, and operational knowledge managers.

FA57s provide expertise in planning and executing experiments, events, and exercises supported by M&S. FA57s build holistic training environments in support of the commander to ensure every aspect of the training runs effectively:

- FA57 officers enhance soldier and unit readiness and combat effectiveness with the use of live, virtual, constructive simulations, and gaming technology which enables them to create realistic environments of both current and future battlefields.
- FA57s provide operational relevance and technical knowledge to support Army operations, testing, experimentation, combat and material developments.
- FA57s showcase their distinctive abilities by employing technology in support of mission preparation and Mission Rehearsal Exercises (MRE/MRX) to enhance training while optimizing resources to support deploying forces. This results in increased training and realism that is key to success on today's complex and rapidly changing battlefield.

Some examples of FA57 success are evident in the live simulation support to operations at the National Training Center (NTC) at Fort Irwin, CA and the Joint Readiness Training Center (JRTC) at Fort Polk, LA. FA57s also support unit collective training by leveraging virtual simulators such as the Close Combat Tactical Trainer (CCTT) to improve warfighting skills. Through the use of constructive simulations such as One-Semi Automated Forces (OneSAF) and the Joint Conflict and Tactical Simulation (JCATS), FA57s further new concepts and experimentation, as well as provide support to multi-echelon staff training. FA57s frequently perform as Battle Command and Operational

Knowledge Management officers at units Army-wide, improving process and understanding by exploiting information technology and linking it to the unit's warfighting function.

CAREER PROGRAM 36 (CP36)

CP36 [3] is a multi-faceted Army civilian career program for training, educating and developing analysis, modeling and simulation civilian human capital. Developed and run by the Army Simulation Proponent and School, the program provides career guidance and funding enabling individuals to gain the necessary expertise to utilize the full capabilities of analysis, modeling and simulation. Civilians in over 20 different federal job series learn to develop, use, manage and integrate these technologies and processes at all levels throughout the Army and for all functions. Analysis, modeling and simulation is pervasive throughout the Army, and is found in the Acquisition, Analysis, Operations, Testing, Training, Experimentation and Intelligence communities.

CP36 education, training and development are offered in a variety of formats, from certificate courses, industry/university courses, distance learning, resident courses, and rotational assignments. Both FA57 and CP36 offer advanced M&S degree opportunities. In addition, college graduates can enter into a CP36 two-year civilian Internship program.

Civilians in the CP36 program learn to use, apply, and manage analysis, modeling, and simulation to enhance the acquisition process and to analyze Army missions, activities and warfighting capabilities. Those who have experienced CP36 education, development and training often serve as technical experts in support of units, organizations and

commands, providing both the organizations and the Army with improved acquisition, analysis, training, operations and plans, testing, experimentation, and intelligence activities. Specifically, Career Program 36 enables the workforce with:

- Multi-disciplinary knowledge, skills, abilities and experiences
- Training and education in M&S theory, models, tools and analysis
- Understanding of how to integrate models and tools into training, acquisition, analysis & experimentation
- The ability to perform as an agent of change who can infuse M&S capabilities throughout the Army of today and the future
- Exposure to all M&S communities and how M&S are used within the Army

M&S are used throughout the Army, and the CP36 program supports these applications by enabling experts to develop new models and simulations, to develop and review M&S policy, guidance and directives, and to incorporate real-world data.

ARMY MODELING AND SIMULATION SCHOOL

The Army Modeling and Simulation School [4] was established in 2005 and provides qualification & certification training for FA57s and education and training for CP36s/other military/civilians. We have continued to educate, train, and develop exceptional modeling, simulation, and battle command professionals with course offerings such as the Simulation Operations Course (6 week resident), the Advanced Simulation Course (2 week resident), the Battle Command Officer Integration Course (2 week resident), the Simulation Operations Professional Course (1-3 week resident/mobile training team), the Simulation (S7) Course (distance learning), the Modeling and Simulation Basics Course (distance learning), the Army Battle Command System & Knowledge Management Synchronizing Course (distance learning), and the Simulation Operations Right Seat Ride (National Training Center 1 week resident).

In summary, the Army continues to lead the way in developing both state of the art modeling, simulation, and system capabilities while also developing the dedicated personnel to effectively utilize them to their greatest potential. Both the FA57 and the CP36 programs are designed to make a difference in soldiers' lives. They create environments that enable knowledge creation and enable the communities through the use of M&S tools and systems. A major benefit seen from these programs continues to be the bridging of the gap between the technology and the warfighters. The growth of the military and civilian programs continues to reflect the Army's sound investment in the military and civilian professionals to meet the most pressing challenges of today's operational, technical, and resource constrained environments. The enduring education, assignments, and technical training not only develop and sustain this essential Army modeling and simulation workforce, but also helps to further our collective knowledge and capabilities across all the Services, DoD and Industry. This is essential as we educate, prepare, sustain and invest in our workforce to advance in today's ever-changing technical world.

REFERENCES

[1] Army Modeling and Simulation, Office of the Deputy Chief of Staff, G-8/CAA at **http://www.ms.army.mil/index.html .**

[2] FA57 Simulation Operations Officer, Office of the Deputy Chief of Staff, G-8/CAA at **http://www.ms.army.mil/FA57/index.html.**

[3] CP36 Analysis, Modeling & Simulation, Office of the Deputy Chief of Staff, G-8/CAA at **http://www.ms.army.mil/cp36/index.html.**

[4] Army Modeling and Simulation School, Office of the Deputy Chief of Staff, G-8/CAA at **http://www.ms.army.mil/school/index.html.**

Training Platoon Leader Adaptive Thinking Skills in a Classroom Setting

AUTHORS

Jason Sidman
Aptima, Inc.
Woburn, MA
sidman@aptima.com

Robert J. Pleban
Army Research Institute
Fort Benning, GA
robert.pleban@us.army.mil

Alexandra Geyer
Aptima, Inc.
Woburn, MA
ageyer@aptima.com

Rob Semmens
Imprimis, Inc.
Colorado Springs, CO
rob.semmens@imprimis-inc.com

ABSTRACT

Given the irregular and unconventional nature of current military conflicts, a major objective for military training and education is to develop "highly adaptable leaders that can quickly hone unit skills on an assigned mission, can reach back to leverage sources of expertise before and during mission execution, can rapidly adjust to changing circumstances, and can aggressively learn from previous and current operations" (AR 350-1, 2007). Today's Soldiers must be able to think critically, make rapid and accurate decisions, and solve complex problems. However, to develop instruction that is designed to train such cognitive skills may require the use of training approaches that are currently either not employed at all or employed very rarely within traditional military education. This paper reports on the results of an effort examining the development and evaluation of an exemplar training module designed to train adaptive thinking in the context of troop leading procedures (TLP), and that can be used within the real circumstances and constraints of a typical military educational environment. The training is based on constructivist principles of experiential learning and draws heavily from approaches such as contrasting cases/invention (e.g., Schwartz & Bransford, 1998). The approach requires students to exercise adaptive thinking skills in response to changing conditions during mission planning that have been engineered to contrast with previous conditions in order to demonstrate important principles of planning (e.g., terrain-based planning vs. enemy-based planning). Forty-two participants from the Infantry Basic Officer Leader Course (IBOLC) participated in the pilot trial of the new approach. Practical implications for adopting this training methodology within Army institutional training are identified.

ABOUT THE AUTHORS

DR. JASON SIDMAN is the Leader of the Multimedia and Instructional Design team at Aptima. His work at Aptima focuses on the design of pedagogical approaches and the development of computer-based training systems.

DR. ROBERT J. PLEBAN is a research psychologist at the U.S. Army Research Institute for the Behavioral and Social Sciences at Fort Benning, Georgia. His research has examined innovative ways of training critical thinking skills in the areas of ethical decision-making/adaptability and enhancing transfer of learning.

DR. ALEXANDRA GEYER is a Cognitive Scientist at Aptima. Dr. Geyer's expertise is in the areas of psychology, neurophysiology, cognitive processing and adaptive training system design.

MR. ROB SEMMENS is a decorated combat infantryman who has studied training needs and the technology required to meet those needs.

BACKGROUND

The demands facing small unit leaders (platoon, squad, team) in the contemporary operational environment (COE) require that they demonstrate a high level of adaptability. Leaders must be able to adjust rapidly to new and unforeseen circumstances across a wide variety of operations

including humanitarian assistance, peacekeeping, peace enforcement, and low intensity conflict as part of a joint, combined, or interagency operation (TRADOC PAM 525-66, 2001). High operational tempo, increased uncertainty, cultural differences, a determined and resourceful enemy, and the need to constantly shift tactics and approaches are some of the key factors which have contributed to an environment where adaptability is required for mission success (Mueller-Hanson, White, Dorsey, & Pulakos, 2005). The Army, more than ever, needs "… agile and adaptive leaders able to handle the challenges of full spectrum operations in an era of persistent conflict." (FM 3-0, 1-83, 2008).

Adaptability has been defined in many ways (e.g., Pulakos, Arad, Donovan, & Plamondon, 2000; Smith, Ford, & Kozlowski, 1997). The definition adopted for this research is the one provided by Mueller-Hanson et al. (2005) who define adaptability as an effective change in response to an altered situation. Underlying this definition is the notion that, for an individual to respond in an adaptive fashion, he or she must first recognize the need to change based on some perceived alteration in the environment and then change his or her behavior in an appropriate manner.

Objective: Training Adaptability

Institutional courses such as the Infantry Basic Officer Leader Course (IBOLC) are tasked with providing new lieutenants with the fundamental knowledge and skills that will enable them to function effectively as platoon leaders in their first unit of assignment. Not surprisingly, the operational needs of units have impacted course content. In addition, the need to rapidly fill platoon leader positions in operational units may shape how topic areas are taught, which will limit how such content domains as adaptability are addressed in these (institutional) settings.

Thus, only select adaptability attributes or characteristics may be amenable to training at the institutional level. Attributes such as personality and cognitive ability, while predictive of adaptive performance, would be less amenable to training interventions and have a low payoff with regard to improved adaptive performance relative to the costs of developing training in these areas. On the other hand, attributes such as domain specific knowledge, (varied) experience, and, to a lesser extent, metacognition, and problem solving skills are much more amenable to training within an institutional setting (Mueller-Hanson et al., 2005).

Institutional training is typically formal and structured, involving both classroom training and field training in a controlled environment. The focus of this research is on designing effective and efficient classroom training to enhance the adaptive/critical thinking process, i.e., to provide the basic knowledge, concepts, and skills that will lay the ground work for future learning and, importantly, will enhance the transfer of knowledge to novel situations (a key component of adaptability). More specifically, this research will examine adaptability/critical thinking as applied to the mission planning and analysis process. This is a very challenging task for junior leaders. The fast paced, rapidly changing nature of operational missions requires that the platoon leader be able to quickly assess situations, identify key aspects of the planning process, and create follow-on orders which reflect an awareness of these factors (i.e., the changing situation and its impact on earlier plans).

OVERVIEW OF TRAINING STRATEGIES

Three general learning strategies were considered to guide the development of the mission planning module and are briefly described below.

Inquiry Based Learning (IBL)

Inquiry or problem based-learning is founded on research which suggests that by having students learn through problem solving experiences, they can learn both content as well as thinking strategies. In IBL, students learn through facilitated problem solving. More specifically, learning centers on a complex problem that does not have a single correct answer. Students work in collaborative groups to identify what they need to learn to solve a problem. They engage in self-directed learning and then apply their new knowledge to the problem. They then reflect on what they learned and the effectiveness of the strategies employed. In this approach, the instructor's role is to facilitate the learning process rather than provide knowledge. Because students are self-directed, managing their learning goals and strategies to solve ill defined problems, they are able to, presumably, acquire the skills needed for lifelong learning (Hmelo-Silver, 2004; see also Hmelo-Silver, Duncan, & Chinn, 2007).

Guided Experiential Learning (GEL)

The GEL approach to learning is based on a large body of research which indicates that providing information does not equate to training. Furthermore, under the GEL model, providing trainees with a field-based problem or an immersive situation alone are not adequate to achieve individual or team learning (Mayer, 2004). A GEL-based course module is grounded on the premise that strong early guidance for the learning of expert-based strategies for task performance works best.

Guidance consists of clear procedures, accurate demonstrations of authentic field-based problem solving, and practice on increasingly difficult problems where expert feedback helps correct trainee misconceptions concerning the correct performance of the task. Guidance is gradually faded until the trainee is able to continue to learn and perform at or above expectations (Clark, 2004).

The structure of a GEL lesson or module follows the same format regardless of the problem. Typically, lessons are sequenced in the following order. The lesson starts with the instructor presenting students with a learning objective (to give the trainees an end state), then telling them why (to motivate learning) and what will happen in the lesson (an overview) to create a mental model of what will be learned. The instructor then teaches the conceptual knowledge needed to learn the procedure (if any), demonstrates the procedure and provides practice and feedback (Clark, 2004).

The overall quality of a GEL lesson is a direct function of the cognitive task analysis (CTA) that is performed in the course design phase. A CTA is a knowledge elicitation procedure designed to uncover information about the knowledge, thought processes and goal structures that underlie observable task performance (Clark, Feldon, van Merrienboer, Yates, & Early, 2007). Execution of Clark's CTA approach is highly structured (Expert Knowledge Solutions, 2007) and requires extensive training (and certification) of the interviewer before he/she is permitted to conduct a CTA (Clark's version) without supervision.

Not all courses are candidates for GEL design. Courses for advanced learners and or/experts do not require the learning support provided in a GEL designed course. In general, when the learning goals of a course are vague or the problems addressed in the course are unstructured/ill defined, and when only conceptual knowledge is being taught (i.e., without "how to" instruction) GEL design is not useful (Clark, 2004).

Contrasting Cases/Invention

Contrasting cases/invention are two instructional design features used to enhance deep understanding of subject matter materials. The approach was developed to help learners construct new knowledge for themselves and become more adaptive/effective problem solvers (Schwartz & Bransford, 1998; Schwartz & Martin, 2004). A key objective of this approach is to optimize the use of lectures/reading text materials to develop these skills. Schwartz and Bransford argue that the value of lectures can be enhanced if the trainee is able to map information from the lecture or text into the knowledge of the problem situation that they have already developed as a result of their prior experiences. A key assumption of this strategy is that the trainee can activate the prior knowledge. Schwartz and Bransford propose a way for activating this prior knowledge through the use of contrasting cases/invention. Based on theories of perceptual learning that emphasize differentiation (e.g., Bransford, Franks, Vye, & Sherwood, 1989), providing trainees with opportunities to analyze sets of contrasting cases (e.g., analyzing the results from different experiments, key aspects of different theoretical models) can help them become sensitive to information that they might not otherwise notice. Contrasting cases help attune people to specific features and dimensions that make the cases distinctive. The refined information provides the foundation for guiding other activities such as creating images, elaborating, and generating questions, which can enhance development of adaptive problem solving skills.

According to Schwartz and Martin (2004), contrasting cases can help learners pick up or notice distinctive features of a problem; however, it is their actions that are critical for helping them discern the deep-level structures that organize those features. To make contrasting cases effective, learners need to undertake productive activities that lead them to notice and account for contrasts in the different cases. Schwartz and Martin use the term "invention" to describe this process. Invention involves production activities, like inventing solutions that can be particularly beneficial for developing early knowledge and facilitating learning. These solutions could, for example, be in the form of graphs, or general formulas. Invention can help develop and/or clarify interpretations of the problem in question by forcing students to notice inconsistencies in

their approach or mental model of their solution and work to reconcile them. This, in turn, provides the knowledge that will prepare them to learn from subsequent instruction (lectures) with deeper understanding (Schwartz, Sears, & Chang, 2008).

As with IBL, to optimize deep understanding of the subject matter material, Schwartz and colleagues advocate a particular sequencing of events. Students first try to solve novel problems without guidance/instruction. Then, they receive direct instruction and demonstrations regarding the tasks. Finally, they apply what they have learned to novel situations. For example, students might analyze data sets from classical experiments and attempt to graphically display the general phenomena from the data. Or, they might be asked to invent a model or formula that will accurately describe the concept (e.g., reliability or correlation). This would be followed by a lecture and (sometimes) class discussion. Finally, students would be presented with new problems and asked to make predictions concerning the outcomes of new experiments or applying the formula or model to solve another (novel) problem (Schwartz & Martin, 2004; Schwartz, Bransford, & Sears, 2005).

METHODOLOGY AND DESIGN OF A HYBRID TRAINING APPROACH

The training approach developed for this effort combined elements of IBL and Invention. Considerations for not using the GEL approach include the extensive time involved in training personnel to conduct and accurately execute a CTA, as well as the time involved to train instructors in the GEL approach, and the inability of GEL to address the key objective of the proposed training module - to develop the conceptual skills (adaptive/critical thinking) needed to produce effective solutions (plans) which have no clearly defined right or wrong answer. Because the goal was to promote adaptive thinking, we believed that these two approaches were more appropriate than GEL. Prior research suggest that direct instruction (e.g., GEL) is very effective in training procedural skills and the acquisition of facts, while constructivist (e.g., IBL, Invention) approaches are more effective in promoting cognitive skills like adaptability (e.g., Duffy and Kirkley, 2007). For all of these reasons, our hybrid approach therefore combined elements of both IBL and Invention.

The IBL Influence: Sequence of Activities

While there are variations on the IBL approach, the current training strategy requires the participants to work on multiple exercises (missions) prior to receiving any lecture or extensive discussion; a distinguishing characteristic of the IBL approach. Following the lecture, participants are then presented with another mission, related to the earlier ones (for additional practice). Finally, the participants receive a very different mission to assess near transfer (i.e., whether the newly acquired knowledge is successfully applied (transferred) to a novel problem/situation). Thus, while the design factors are the same in both the current approach and the Mueller-Hanson approach, (i.e., lecture, multiple exercises [or exposure to multiple examples], discussion/feedback), the key difference between the approaches is the sequencing of activities. By beginning with a problem, IBL advocates argue, the learner becomes more prepared to learn from the lecture. As they argue, there is a "time for telling" (e.g., Schwartz & Martin, 2004).

The Invention Influence: Contrasting Cases

While the sequencing of activities (problem before lecture) represents a framework for the training events, the selection of what those events should cover is a critical instructional consideration; this is where Contrasting Cases influenced the current approach. Given the goal of promoting adaptive thinking, the multiple exercises need to not only differ from one another, but differ in a meaningful way. Indeed, the power of designing such "contrasting cases" is that the student discovers the desired instructional outcome (i.e., the dimension along which two cases contrast). The link between contrasts and training objectives distinguishes this approach from general "what-if" exercises (though these can certainly help trainees consider contingencies at a general level).

In the present context, the desired educational outcome was for students to understand the dynamic relationship of friendly, enemy, and terrain components of terrain analysis when developing operational orders (OPORDs). New lieutenants might treat each of these components in an isolated, static fashion because they are focused on writing the OPORD rather than understanding the mission. The contrasts were therefore designed to demonstrate to the student that changes to any one component (friendly, enemy, terrain) will affect the other two. The contrasts are described in more detail below.

In the first scenario, the company OPORD described the plan for an offensive operation. The company mission was to clear Objective (OBJ) Anvil, and the third platoon's mission was to secure a mosque, which would enable the company main effort, second platoon, to clear the rest of OBJ Anvil. Included in the company OPORD were the area of operations/interest, situation (enemy and friendly), terrain and weather, concept of operations, attachments and detachments, company mission, commander's intent, tasks to maneuver units, and coordinating instructions.

The first Fragmentary Order (FRAGO) changed the task of third platoon from "secure" to "isolate", thereby changing the entire operation from being focused on the terrain (the bomb making facility) to being focused on the enemy (bomb making expertise). Table 1 summarizes the changes in the OPORD and their intended impact on the participant's (platoon leader's) analysis/development of his OPORD.

The new lieutenant may not fully appreciate the power of the meanings of the tactical mission tasks (secure vs. isolate). Consequently, s/he may simply change the actual words in the revised OPORD rather than changing the plan conceptually. However, what they should come to realize, and what the instructor should help them discover in the lecture following FRAGO 1, is that the change to the friendly mission changes how they should analyze the enemy and the terrain; indeed, terrain analysis is dynamic.

In FRAGO 2, a high value target (HVT) is said to be on the objective. Again, the new lieutenant could simply add these words to the OPORD but keep the plan relatively unchanged. However, having worked through the OPORD and FRAGO 1, and having received the lecture following FRAGO 1, s/he should consider the dynamic nature of the analysis and consider how this knowledge of the enemy will affect friendly forces and the terrain. For example, s/he should expect the enemy now to fight in order to allow the HVT to escape. S/he may not have been prepared for that possibility before.

OPORD	FRAGO 1
Platoon's primary task is to "secure"	Primary task changes to "isolate"
Battalion operation is terrain focused	Battalion operation becomes enemy focused

Table 1. Example of a Contrasting Case

LESSONS LEARNED

The project team evaluated this training approach in a two-day pilot session with 42 male second lieutenants who recently graduated from the Infantry Basic Officer Leader Course (IBOLC). Demographics are presented in Table 2. Because we were evaluating a new training approach, we present lessons learned from all aspects of the project: development of materials, execution of the classroom session, and analysis of findings. But we first start with a more detailed description of the procedure.

AGE	YEARS IN MILITARY	PRIOR ENLISTED	DEPLOYED IN OIF/OEF
M= 23.4	M = 4.0	7 (16.7%)	4 (9.5%)
SD = 1.8	SD = 1.9		

Table 2. Participant Demographics

Procedure

The participants were provided with notebooks and different colored pens and instructed to do all their work, except graphics and concept sketches in the notebooks. The instructor role played the company commander and gave the area of operations (AO) orientation briefing and company OPORD. The Area of Operations brief was similar to a briefing a unit might get during a Relief in Place/ Transfer of Authority (RIP/TOA), and while not entirely doctrinally correct, provided the appropriate background information to allow students to familiarize themselves with the situation. For assessment purposes, the participants were asked to record any questions they had for the company commander in their notebooks.

Each participant role played a platoon leader for 3rd platoon, Alpha company, and was asked to write their own individual platoon order. They were allowed to use whatever OPORD format they wanted (e.g., matrix). If they felt constrained by time they were instructed to focus on what was important, just as they would do in a unit.

The participants then began work on their backbriefs and Warning Order (WARNO). When they finished, they were instructed to start on the OPORD. The students were allotted a total of two hours (with an hour break for lunch) to complete the backbrief, WARNO, and OPORD before they received the first FRAGO (FRAGO 1). They then began to revise their OPORD based on FRAGO 1 by making changes to their base plan (using a different color pen).

After participants received FRAGO 1 and worked for 45 minutes, the instructor provided a lecture. The focus of the lecture was to emphasize the overall importance of developing a model of the plan and mentally playing out the plan (mental simulation). In addition, the instructor discussed how the changes in FRAGO 1 differed from the original company OPORD (part of the contrasting case strategy). The goal was to highlight distinctive features in the two plans (original OPORD and FRAGO 1), e.g., implications between isolate and secure, presence of high value target (experienced IED maker), changes in the battalion focus (neutralize, contain, and defeat) and how that impacted FRAGO 1. The instructor closed the lecture by asking the participants what they would add/change to their OPORD based on the changes noted on FRAGO 1, and how would these changes show up on their platoon OPORD.

The instructor then passed out FRAGO 2 along with a different colored pen in order to track the changes made in each phase. The participants were then given time to update their order. When the participants completed FRAGO 2, the instructor conducted a brief discussion designed to highlight second order effects (e.g., Did you do the mission at night with night vision goggles or with white light? If you used night vision goggles (NVGs), how did you account for the Iraqi squad that probably did not have NVGs?).

Following the discussion, the participants received the second scenario (transfer task) which was very different from the first OPORD. The transfer task was a stability operation (secure a market place). In contrast, the first OPORD and follow on FRAGOs were part of an offensive operation. The objective was to determine how well information provided in the contrasting cases and lecture and employed in FRAGO 2 generalized (transferred) to the more nebulous stability operation. For example, we intended to see if students considered the actions of the enemy after they had secured the market—how would they attempt to further disrupt the market given a new security posture? How would they neutralize the terrain features that most affected the marketplace? How would they incorporate other combat multipliers for full advantage, such as the engineers or civil affairs team?

Lessons Learned: Course Development

The development of course materials that promote adaptability is a challenging, but potentially liberating, task. One approach to training adaptability would be to create a course about adaptability and associated constructs and concepts. The approach we endorse here, however, has a subtle but important feature; adaptability is trained in the context of the existing course curriculum. Terrain analysis was still covered, as it is in the existing IBOLC program of instruction (POI). But, it was covered in a way that not only teaches students about terrain analysis, but promotes adaptive thinking at the same time. This value added is the great potential benefit of this approach.

Designing the contrasts, however, is a challenging task. As described earlier, though both contrasting cases and what-if exercises promote contingency planning skills, contrasts are meant to be more illustrative of training objectives than traditional what-if exercises. But they require extensive effort and thought to design, and it can be difficult to know that the students will discover

the same underlying dimension of the contrast that the instructor sees. This can, of course, be mitigated to an extent by the keen instructor-guided facilitation of discussion.

While contrasting cases/invention is a critical part of Schwartz's approach, the lecture component is equally valuable. It offers a higher level explanation of the concept/phenomena that would be quite difficult and time consuming for the student to discover on his or her own. The higher level explanation is important because it provides a generative framework that can extend one's understanding beyond the specific cases that have been analyzed and experienced (Schwartz & Black, 1996) and thus, enhances adaptive problem solving in general (transfer). Schwartz, Bransford, and Sears (2005) present evidence that the most effective design combination includes both opportunities for invention and analysis (contrasting cases) followed by opportunities for learning efficient solutions derived by experts (typically) presented in lecture format.

Lessons Learned: Procedure

While a constructivist approach is appealing from the standpoint of developing problem solving skills which may be applied to similar situations outside the initial training environment, there were several drawbacks to this strategy for the current research. For example, issues involving classroom organization (shorter instructional periods in IBOLC with often strict time constraints), skill levels of current instructors to serve as course facilitators for this approach, and the relatively high IBOLC student/instructor ratio (40:1) threaten the practicality of a constructivist approach. For example, it became immediately clear during the execution of the session that an extremely long amount of time elapsed between the AO brief and when students received any feedback (following FRAGO 1). This was intentional; we wanted students to get deep enough into the problem, and develop a strong enough commitment to a plan in response to the OPORD, that the introduction of a change (FRAGO 1) would significantly impact them. However, the theory behind the sequencing was diminished by the practice of the sequencing; participants appeared fatigued after working independently all morning on their OPORD. Despite being given less time to work on FRAGO 1, in some sense the damage to their motivation had been done. The participants seemed much less able to commit their full attention to FRAGO 2, and even less to the transfer task.

In addition, instructor selection and training would therefore be heavily impacted by a major commitment to adopting approaches such as these in Army institutional training. Ideally, the instructors themselves would be adaptive thinkers, capable of and comfortable with deviating from the course plan in order to facilitate classroom discussion. However, current training courses for instructors typically do not address such skills.

Lessons Learned: Evaluation

Initially the team had planned to have experts rate participants' adaptive thinking on the OPORD, FRAGO 1, and FRAGO 2 using a set of behaviorally anchored ratings scales (BARS) developed by Phillips, Ross, and Shadrick (2006). The logic was that participants would demonstrate greater adaptive thinking following FRAGO 1 than preceding it (by virtue of having experienced the contrasting cases and the lecture). However, this proved a troublesome method. For one, we conducted the training in a single day. The expectation that participants would become adaptive thinkers after a single day of instruction was unrealistic; the measure was therefore not going to be sensitive enough. In addition, the effort to score many sets of OPORDs is great as well (in fact, the grading of the OPORDs continues as of the writing of this paper). An instructor would struggle to provide timely feedback to students based on their evolving responses to three versions of an OPORD.

Assessment in general can be challenging in a constructivist approach. For one, rather than cumulative assessment (i.e., a grade at the end of an exercise), constructivists emphasize formative assessment (i.e., feedback as part of the exercise). This presents another institutional challenge, as it complicates the award of promotions and other recognition currently based on grades. Second, grading thinking skills as opposed to procedural skills may also be new and unfamiliar to instructors. Indeed they need to understand at a deep level how the outputs of procedural skills are connected to thinking skills. The anchors of the BARS were intended to help make this connection, and a replacement tool would require the same connection. Such a tool would require extensive design and validation, placing yet another burden on the institution.

CONCLUSIONS

Constructivist theorists provide compelling reasons to employ their methods for training adaptive thinking. Similarly, experiences with using direct instruction to train such cognitive skills can be unsatisfying (not just in the military of course, as almost all of us can attest to). Consequently, there seems to be a willingness and an openness to adopting new methods of instruction to train such skills (e.g., soon to be released Army training manuals will explain that different training approaches are appropriate for training different types of skills).

However, the institutional barriers to incorporating new approaches are tremendous. As described earlier, class sizes and schedules alone make the adoption of constructivist approaches prohibitive. Furthermore, changing the way instructors are trained to do training (to be facilitator rather than conveyor) would be a massive undertaking as well.

Constructivist approaches explored in this research were not successful due, in part, to some of the institutional training constraints identified earlier in this paper (not unique to only the military). This presents a challenge to constructivist theorists: how can approaches be implemented in this training environment given these constraints?

Indeed, that was the question we aimed to answer in this effort, and we believe we have identified potential parts of the solution, as well as additional constraints. For example, while the use of working through contrasting cases in the context of an actual operations order exercise is appealing and, we continue to believe, pedagogically valuable, a more targeted task, or subtask, could address some of the time and fatigue pressures experienced during our exercise. Perhaps focusing simply on developing concept sketches, for example, would have required the same kind of thinking but with less of the cognitively tangential tasks. Or perhaps eliminating the backbrief and WARNO and focusing more on the OPORD would have saved time and effort. However, part of the reason why we did not do this ahead of time is that we were unable to find any descriptive guidance on how to develop constructivist approaches. The guidance we did find seemed vague.

Finally, training cognitive skills takes more than one day. It will almost certainly require repeated exercises over several classroom sessions with follow-on lectures and discussion to highlight key learning points and insure deeper understanding of the concepts presented. We were constrained logistically to one day, but instructors too would have to plan for several such exercises rather than a single one.

In conclusion, constructivist approaches hold a lot of promise for training the cognitive skills essential in the operational environment; however they require significant engineering to be implemented in the institutional environment.

ACKNOWLEDGEMENTS

We would like to thank the Infantry Basic Officer Leader Course (IBOLC) at Fort Benning for their support of this project. The views, opinions, and/or findings contained in this report are those of the authors and shall not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documents.

REFERENCES

Bransford, J. D., Franks, J. J., Vye, N. J., & Sherwood,R. D. (1989). New approaches to instruction:Because wisdom can't be told. In S. Vosniadou & A.Ortony (Eds.), Similarity and analogical reasoning(pp. 470-497). New York: Cambridge UniversityPress.

Clark, R.E. (2004). Design document for a guidedexperiential learning course. Los Angeles, CA:Institute for Creative Technologies.

Clark, R. E., Feldon, D. F., van Merrienboer, J. J. G.,Yates, K. A., & Early, S. (2007). Cognitive taskanalysis. In J. M. Spector, M. D. Merrill, J. J. G. vanMerrienboer, & M. P. Driscoll (Eds.) Handbook ofresearch on educational communications andtechnology (3rd ed.). Mawah, NJ: Erlbaum.

Duffy, T.M., & Kirkely, J.R. (2007). Experientiallearning. Presentation at Redstone Arsenal,Huntsville, AL.

Expert Knowledge Solutions. (2007). Job aid forcompleting a CTA report. Available from the firstauthor.

Hmelo-Silver, C. E., (2004). Problem-based learning:What and how do students learn? EducationalPsychology Review, 16(3), 235-266.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A.(2007). Scaffolding and achievement in problembasedand inquiry learning: A response toKirschner, Sweller, and Clark (2006). EducationalPsychologist, 42(2), 99-107.

Mayer, R. E. (2004). Should there be a three-strikesrule against pure discovery learning? The case forguided methods of instruction. AmericanPsychologist, 59(1), 14-19.

Mueller-Hanson, R. A., White, S. S., Dorsey, D. W., &Pulakos, E. D. (2005). Training adaptable leaders:Lessons from research and practice (ResearchReport 1844). Arlington, VA: U. S. Army ResearchInstitute for the Behavioral and Social Sciences.

Phillips, J.K., Ross, K.G., & Shadrick, S.B. Users Guidefor Tactical Thinking Behaviorally Anchored RatingScales. (Research Product 2006-05). Arlington, VA:U. S. Army Research Institute for the Behavioral andSocial Sciences.

Pulakos, E. D., Arad, S., Donovan, M. A., &Plamondon, K. E. (2000). Adaptability in theworkplace: Development of a taxonomy of adaptiveperformance. Journal of Applied Psychology, 85(4),612-624.

Schwartz, D. L., & Black, J. B. (1996). Shuttlingbetween depictive models and abstract rules:Induction and fallback. Cognitive Science: AMultidisciplinary Journal, 20(4), 457-497.

Schwartz, D. L., & Bransford, J. D. (1998). A time fortelling. Cognition and Instruction, 16(4), 475-522.

Schwartz, D. L., Bransford, J. B., & Sears, D. (2005).Efficiency and innovation in transfer. In Transfer oflearning from a modern multidisciplinary perspective(pp. 1-51). Information Age Publishing.

Schwartz, D. L., & Martin, T. (2004). Inventing toprepare for future learning: The hidden efficiency ofencouraging original student production in statisticsinstruction. Cognition and Instruction, 22(2), 129-184.

Schwartz, D. L., Sears, D., & Chang, J. (2008).Reconsidering prior knowledge. Retrieved from **http://aaalab.stanford.edu/papers/Schwartz-Reconsidering-Prior-K.pdf**

Smith, E. M. Ford, J. K., Kozlowski, S. W. (1997).Building adaptive expertise: Implications fortraining design strategies. In M. A. Quinones & A.Ehrenstein (Eds.), Training for a rapidly changingworkplace (pp. 89-118). Washington DC: AmericanPsychological Association.

U. S. Department of the Army (2001). Objective forcecapability (TRADOC Pamphlet 525-66).Washington, DC: Author.

U. S. Department of the Army (2008). Operations. Washington, DC: Author.

This paper was adopted from I/ITSEC 2010 and is reprinted with permission.

A Model for University-Level Education in Modeling and Simulation

AUTHORS

Roland R. Mielke
James F. Leathrum, Jr.
Frederic D. McKenzie
Department of Modeling, Simulation and Visualization Engineering
Old Dominion University

ABSTRACT

There are at least two student constituencies for university-level modeling and simulation (M&S) educational programs, the user constituency interested in using M&S to study another discipline and the developer constituency interested in studying M&S as a discipline. This paper describes the development of a multi-faceted approach to M&S education. The user student constituency is served at the graduate level by M&S certificate programs implemented across all of the academic colleges and at the undergraduate level by a minor in M&S. The developer constituency is served by an academic department that offers bachelor's, master's and doctoral M&S degree programs. Both student constituencies are further supported by a university-wide research center focused on M&S research and development activities.

KEYWORDS

Higher education, academic programs, modeling and simulation (M&S), program organizational model, curriculum design

1. INTRODUCTION

Over the past decade, the notion that computer simulation rivals in importance the more traditional aspects of scientific investigation, theory and experimentation, has gained wide acceptance. Several recent government-sponsored panels have accentuated these observations and have noted the urgent need to develop and enhance educational programs in simulation. The NSF Blue Ribbon Panel on Simulation-Based Engineering Science (SBES) [1] states that "seldom have so many independent studies by experts from diverse perspectives been in such agreement: computer simulation has and will continue to have an enormous impact on all areas of engineering, scientific discovery, and endeavors to solve major societal problems." Regarding education in computer simulation, the report goes on to state: "The old silo structure of educational institutions has become an antiquated liability. It discourages innovation, limits the critically important exchange of knowledge between core disciplines, and discourages the interdisciplinary research, study, and interaction critical to advances in SBES." The President's Information Technology Advisory Committee (PITAC) Report [2] concludes that "Universities must implement new multidisciplinary programs and organizations that provide rigorous multifaceted education for the growing ranks of computational scientists the nation will need to remain at the forefront of scientific discovery."

Historically, modeling and simulation (M&S) has been viewed as an important research tool in numerous disciplines or application domains. Research in most domains often proceeds through a sequence of phases that include understanding, prediction, and control. In the initial phase, we are interested in understanding how events or objects are related. An understanding of relationships among objects or events then allows us to begin making predictions and ultimately to identify causal mechanisms. Finally, knowledge of causality enables us to exert control over events and objects. Research moves from basic to more applied levels as we progress through these phases. M&S is closely linked to all of these phases. At the basic levels, research is guided heavily by theory. Models are often used to represent specific instances of theories, discriminate among competing theories, or evaluate underlying assumptions. Likewise, simulations are used to test predictions under a variety of conditions or to validate theories against actual conditions. At the applied levels, simulations are also used to control events and objects. One of the primary uses for simulation is training where the goal is to control performance variability by improving operator reliability. Simulations in the form of mock-ups or prototypes are used in the creation of products and systems to validate predictions regarding operational requirements, specifications, and user/customer satisfaction.

Beginning in the mid-1990's, a second type of M&S professional began to emerge. Motivated by the rapidly growing use of simulation for training, analysis, and decision support by industry and government, these individuals are more interested in learning about M&S rather than just using M&S to study something else. Coming from backgrounds in mathematics, computer science, and engineering, these individuals are interested in the fundamental principles and theoretical foundations of M&S. They are anxious to investigate some of the major challenges of M&S: multi-scale and multi-resolution M&S, interoperability of simulations, composability of models, verification and validation, distributed and real-time simulation, and representation of increasingly complex and data-intensive system problems. In short, this group views M&S as a discipline. Their objective is to obtain a formal education in the M&S discipline and then to find employment opportunities as M&S scientists and engineers.

The growth of the view of M&S as a discipline is well documented in the literature. Since the late 1990's, a number of papers have been written stating the importance and urgency for developing educational programs in the discipline of modeling and simulation. These papers identify desirable program outcomes [3], present suggestions for course and curriculum content [4], [5], and describe potential approaches for, and challenges in, implementing a modeling and simulation program [6], [7], [8], [9]. More recently, curricula [10], [11] and models [12] for graduate modeling and simulation programs have been described. Graduate modeling and simulation programs have been started at several universities including the University of Alabama – Huntsville [13], Arizona State University [14], California State University – Chico [15], Georgia Institute of Technology [16], Old Dominion University [17], and the University of Central Florida [18]. At the undergraduate level, several universities have developed tracks or concentrations focusing on narrow sub-areas of modeling and simulation as part of other degree programs. However, to date, no ABET accredited engineering program in modeling and simulation has been fully implemented.

The recognition of the need for university-level academic programs in M&S is recent and still is embraced by only a handful of universities. Even in universities that are developing M&S programs, there are shortcomings. Most existing M&S academic programs have been developed at the master's level. They are focused primarily on providing a technical education in preparation for working in industry. There are very limited opportunities for study at the doctoral level; however, if M&S is to grow as a discipline, there is a need for significant numbers of Ph.D. graduates to serve as M&S faculty. There are no bachelor's level programs to provide the principal workforce needs of a rapidly expanding M&S industry and to provide the feedstock for M&S graduate programs. Little thought has been given to the education of students who wish to use M&S as a tool to learn more about another discipline.

The objective of this paper is to describe the ongoing organizational development of the M&S academic programs at Old Dominion University. The purpose of the paper is two-fold. First, the description may provide a useful model for other universities considering the initiation of M&S programs. Second, it is hoped that the paper will promote additional discussion in the literature concerning the development of M&S academic programs. In Section 2, two M&S student constituencies are described and their different educational needs are identified. Then a new multi-faceted approach to M&S education is described that addresses the needs of both student constituencies. In Sections 3 - 5, the details of several new program components are presented. An M&S graduate certificate program is described in Section 3, the bachelor's program in Modeling and Simulation Engineering is described in Section 4, and the M&S graduate program is described in Section 5. Concluding remarks are stated in Section 6.

2. PROGRAM ORGANIZATION

We assert that there are two primary student constituencies that must be served by M&S academic programs, the user constituency and the developer constituency. The user constituency consists of students who wish to utilize M&S as a tool to investigate another discipline. They need to know enough about M&S to select the best methodologies for their specific problem and then to apply these methodologies in an appropriate way. The developer constituency consists of students who wish to study M&S as a discipline. Their focus is to learn about the technical details of M&S and then to develop new M&S methodologies and to enhance existing M&S technologies. Upon graduation, these students are likely to seek employment in the M&S professional community as scientists, engineers, technical managers, and teachers.

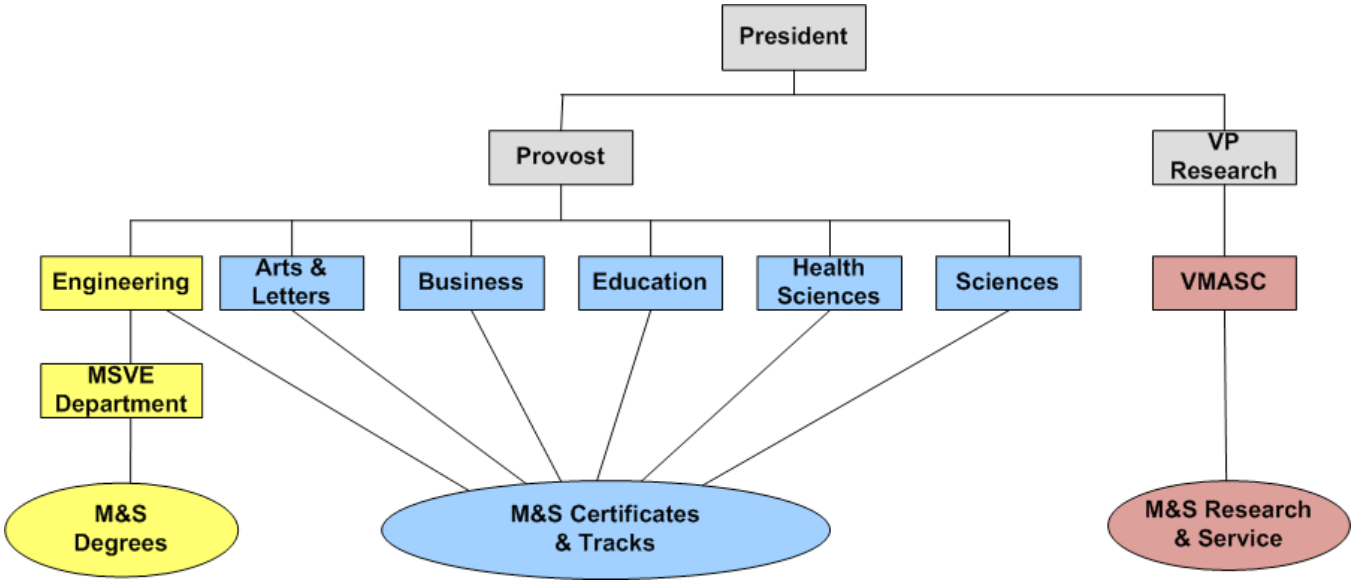


Figure 1. Organization of M&S Programs and Activities

The view of M&S as a discipline leads naturally to the traditional department model for the developer constituency. These students first develop background in mathematics, computer science, and selected engineering topics, and then focus their core technical studies on the M&S Body of Knowledge [19]. They emerge as technical specialists skilled in the design, development, and use of simulation technologies and methods. However, this educational path may not be attractive to the user constituency because of the extensive technical background requirements and the absence of an opportunity to focus deeply on a particular domain area. Thus, it is difficult to serve the needs of all potential M&S students with a "one-size-fits-all" approach. A multi-faceted approach that addresses the needs and requirements of both student types is highly desirable.

The organizational structure of the M&S academic and research programs is shown in Figure 1. There are three important components to this structure. An academic department, the Department of Modeling, Simulation and Visualization Engineering (MSVE), has been established within the College of Engineering and Technology. The MSVE Department administers academic degree programs in M&S at the Bachelor's, Master's, and Doctoral levels, designed to support the needs of the M&S developer student constituency. In addition, the MSVE Department offers an undergraduate minor in M&S, and provides core courses for the graduate certificate programs, both designed to support the needs of the M&S user student constituency. All academic colleges offer graduate certificate programs. The graduate certificate programs consist of clusters of graduate courses designed to enhance the capability to utilize M&S as a tool in other disciplines and result in the award of a certificate of completion. Finally, the university has established a research center focused on M&S research activities. This center, called the Virginia Modeling, Analysis and Simulation Center (VMASC), is administered at the university level through the Office of Research to encourage participation by all academic colleges. It is a place where faculty and students from all disciplines can interact and work on cross-disciplinary projects.

Two university committees, the M&S Steering Committee and the M&S Executive Committee, have been established by the Provost's Office. The M&S Steering Committee consists of M&S faculty representing all six academic colleges. This committee is responsible for recommending policy and procedure and for operational issues spanning all M&S programs. The M&S Executive Committee consists of the dean or associate dean from each academic college. This committee is responsible for approving policy and procedure spanning all M&S academic programs. Together, these committees oversee and coordinate the cross-disciplinary activities for the M&S academic programs.

3. M&S GRADUATE CERTIFICATE PROGRAMS

The M&S graduate certificate programs were designed to support directly the user constituency actively studying a discipline in another department, but requiring M&S skills

for their research. The programs were established to ensure that all faculty and students from across the university have an opportunity to participate in M&S academic programs and research activities. Encouragement to participate was provided by offering a limited number of M&S faculty positions to programs willing to establish a certificate program. Similarly, student interest was encouraged by providing a limited number of M&S graduate assistantships for those students wishing to participate in an M&S certificate program. In a period of only three years, these assistantships have become very competitive. The colleges and departments/programs now offering M&S graduate certificate programs are identified in Table 1.

The requirements for each M&S graduate certificate program consists of four three-credit graduate courses. The selected courses must include the following.

- M&S Fundamentals – coursework offered by the MSVE Department that covers an overview of the M&S Body of Knowledge [19]. Depending on background preparation in mathematics, this requirement can be satisfied with one course or a two-course sequence.
- Models – a course, offered by the certificate host department, that covers models and modeling techniques often used in the certificate discipline.
- Simulation Application – a course, offered by the certificate host department, that demonstrates the application of M&S in the certificate discipline.

A number of the certificate students eventually conduct thesis or dissertation research that utilizes aspects of M&S. Additional support for this research is provided through the staff and facilities of VMASC.

Graduate certificate programs are open to non-degree seeking graduate students as well as degree seeking graduate students. These programs often are attractive to students who desire to enhance their knowledge of M&S for job-related reasons or who are interested in previewing a potential graduate degree program. The university allows up to 12 credits taken in the non-degree seeking status to be applied to a degree program with approval of the administering department.

COLLEGE	DEPARTMENT/PROGRAM
Arts & Letters	International Studies
Business & Public Administration	Decision Sciences & Information Technology
Education	STEM Education
Engineering	Civil & Environmental Engineering Electrical & Computer Engineering Engineering Management & Systems Engineering Mechanical & Aerospace Engineering Modeling & Simulation
Health Sciences	Radiology and Laboratory Sciences
Sciences	Psychology Mathematics & Statistics

Table 1. M&S Graduate Certificate Programs.

4. UNDERGRADUATE PROGRAM IN M&S ENGINEERING (M&SE)

The undergraduate program in M&S Engineering is designed to serve two constituencies, those seeking a major in M&S and those seeking a minor, reflecting the developer and user constituencies respectively. The curriculum is designed primarily to serve those in the major, while recognizing the benefit that a minor in M&S provides to students in other engineering and science fields. This section first presents the goals of the program and then describes the program curriculum.

4.1. Program Goals

The undergraduate M&SE program is designed to meet four sources of program content and goals: the ABET criteria for accrediting engineering programs, the literature defining an M&S body of knowledge, a set of discipline-specific student outcomes identified by program faculty, and university general education requirements. In this section, these requirements are described and the impact on curriculum structure is noted. Discussions of the ABET criteria and the university general education requirements can be found in [20] where the design of the curriculum is presented. The M&S body of knowledge and the discipline-specific student outcomes are briefly discussed here to help set the curriculum in context, though a more complete discussion is found in [20].

M&S Body of Knowledge

A number of significant efforts have been made to define an M&S body of knowledge [4], [21]. However, the focus of these efforts primarily has been to identify content for development of M&S graduate programs or for development of licensure requirements for current M&S practitioners. Thus, while existing body of knowledge presentations may not be entirely appropriate to undergraduate program development, the work serves as a framework from which to select a subset of content areas that are appropriate to M&SE baccalaureate programs. The topics of the body of knowledge displayed here are taken from [19] and were used as the basis for the 2009-2011 revision of the CMSP examination. A more detailed list is found in [21].

1. Fundamental Concepts and Context
2. Applications and Domains
3. Modeling Methods
4. Simulation Implementation
5. Supporting Tools, Techniques, and Resources
6. Business and Management of M&S

Discipline-Specific Essential Knowledge and Skills

The M&SE program should prepare engineering graduates who can utilize modeling and simulation in various domains and for different applications, and who possess the foundation upon which to expand the current M&S body of knowledge. The M&SE program faculty has defined a set of essential knowledge and skills that they believe form the technical foundation for the discipline of modeling and simulation engineering. These are the concepts, principles, and methods that anchor the M&SE curriculum; they represent the fundamentals that every M&SE graduate must know and be able to use. The M&S essential knowledge and skills are stated as a set of student outcomes that are focused on the technical components of the M&S curriculum.

M&SE students who qualify for graduation will have:

1. An ability to communicate designs across technical and non-technical boundaries;
2. An ability to model a variety of systems from different domains;
3. An ability to develop an input data model based on observed data;
4. An ability to select and apply appropriate simulation techniques and tools;
5. An ability to develop simulations in software;
6. An ability to apply the experimental process to acquire desired simulation results;
7. An ability to apply visualization techniques to support the simulation process;
8. An ability to use appropriate techniques to verify and validate models and simulations; and
9. An ability to analyze simulation results to reach an appropriate conclusion.

4.2. Undergraduate M&SE Curriculum

The M&SE curriculum is first and foremost an engineering program having a focus on problem solving, design, and experimentation. The curriculum is designed with 128 credits; 32 credits of mathematics and basic science courses, 58 credits of engineering science and design courses,

32 credits of general education courses, and 6 credits of approved electives. The credit distribution is selected to satisfy ABET Criterion 5 [22] and the university's general education requirements. The curriculum is displayed in "showcase" format in Figure 2. In this display, the courses are distributed over eight semesters and the courses are sequenced to satisfy all prerequisite and co-requisite requirements. The core technical courses are displayed in Figure 3. In this display, the core courses are grouped into three content areas: modeling and simulation; software development; and analysis. This display also shows explicitly all prerequisite and co-requisite relationships among the courses.

The following statements highlight important characteristics of the M&SE curriculum.

- The core engineering science and engineering design courses are grouped into three content areas as highlighted in Figure 3.
- *Modeling and simulation* - The modeling and simulation content area consists of courses that address the design and implementation of models and simulations. This track includes courses on Monte Carlo simulation, discrete event simulation, continuous simulation, and modeling methodologies.
- *Software Development* - The software development content area consists of courses that develop a programming capability and then use that capability to code simulations. Background courses are computer science courses while simulation development is an M&SE course. Computer visualization is included in this content area.
- *Analysis* - The analysis content area consists of the courses that lead to an analysis capability. This area includes the mathematics courses in calculus, differential equations, and probability and statistics. It also includes a simulation analysis course that addresses random number generation, input data modeling, output data analysis, verification and validation, and experimental design.
- A skills content area consists of the courses that provide essential academic skills such as the ability to communicate effectively. The communication skills component is particularly strong and consists of the following course sequence: English composition, technical writing, public speaking, and information literacy and research. Extension and practice of these skills also are present in a number of the core technical courses.

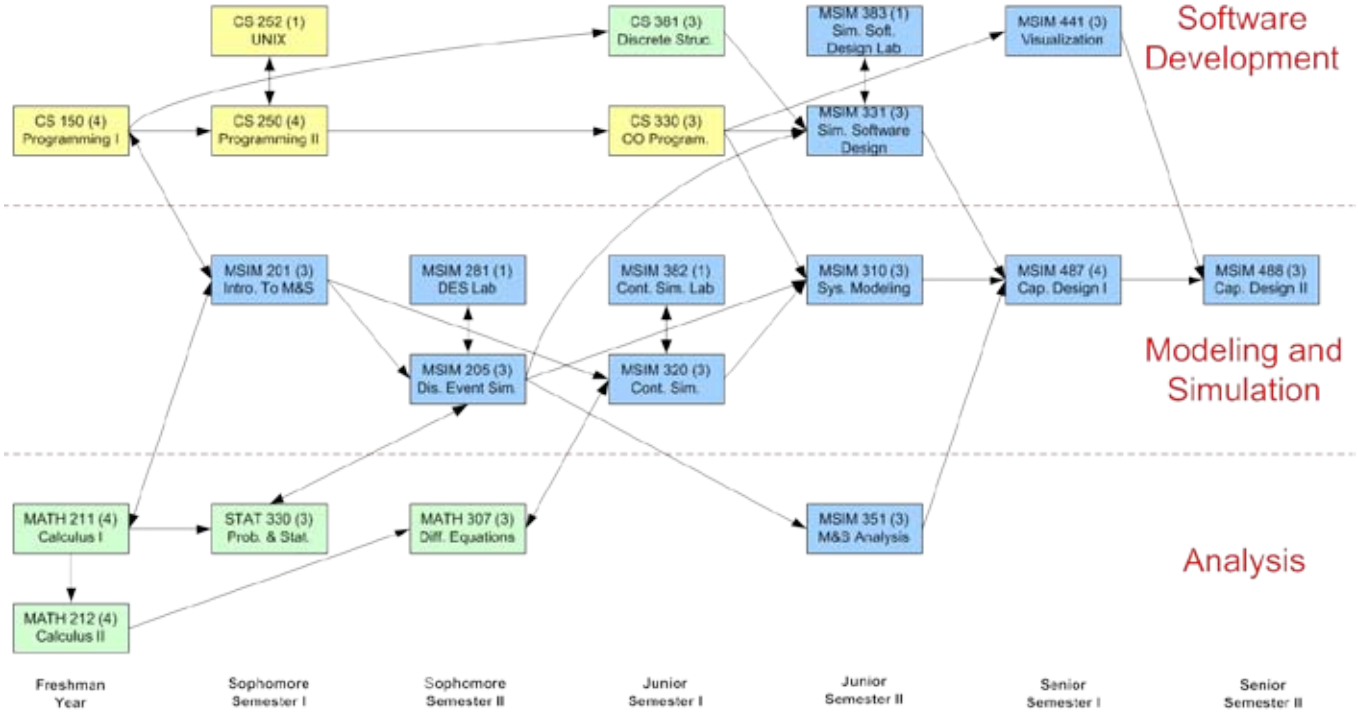


Figure 3. M&SE Core Courses Showing Prerequisite Requirements.

- A professional skills and knowledge content area includes the following courses and topics: introduction to engineering profession; engineering ethics; impact of technology; project management; fundamentals of engineering examination preparation; and the capstone design experience.
- The curriculum allows for inclusion of a formal minor; that is, a set of 12 credits at the upper division offered and approved by an external academic department. We are working with several departments to organize minors that will have a beneficial impact on the student's major studies. For example, the Psychology Department has a minor in human factors that includes courses on perception, cognition, human-computer interfacing, and human behavior modeling. The Biomedical Engineering minor includes courses in biomechanics, human physiology, and innovation and entrepreneurship in the biomedical area. For those students wishing to seek licensure as a professional engineer, minors in civil, electrical, computer, and mechanical engineering are available. The inclusion of a minor in the M&SE curriculum is viewed as a strength and students are encouraged strongly to utilize this option to further round their education and strengthen their potential in the workforce.

4.2. Minor in M&SE

The benefits of getting a solid background in M&S for students in engineering and science fields are becoming more crucial as M&S becomes more and more prevalent in the workforce. Therefore, minor paths through the curriculum have been defined to support educating students having the appropriate mathematics, science and programming skills. As with the overall program, the minor is designed to serve both M&S users and M&S developers.

The M&S user population is considered to include students in all engineering and science fields, as well as others willing to acquire the necessary background. Students in this population usually have a solid background in calculus and differential equations, college level physics, and some basic level programming. Modern computer

Course Number	Course Title	Credits
Freshman Year – First Semester		
MATH 211	Calculus I	4
ENGL 110C	English Composition I	3
CHEM 121N	Chemistry I & Lab	4
ENGN 110	Engineering & Technology I	2
Gen Ed	Oral Communication	3
Freshman Year – Second Semester		
MATH 212	Calculus II	4
CHEM 123N	Chemistry II	3
CS 150	Programming I	4
PHYS 231N	University Physics I	4
MSIM 111	Information Literacy for M&SE	2
Sophomore Year – First Semester		
STAT 330	Probability and Statistics	3
PHYS 232N	University Physics II	4
CS 250	Programming II	4
CS 252	Introduction to UNIX	1
MSIM 201	Introduction to M&S	3
Sophomore Year – Second Semester		
MATH 307	Differential Equations	3
ENGL 231C	Technical Writing	3
Gen Ed	Human Creativity	3
Gen Ed	Literature	3
MSIM 205	Discrete Event Simulation	3
MSIM 281	M&S Laboratory 1	1

Course Number	Course Title	Credits
Junior Year – First Semester		
CS 330	Object-Oriented Prog. & Design	3
CS 381	Discrete Structures	3
MSIM 320	Continuous Simulation	3
MSIM 382	M&S Laboratory 2	1
Gen Ed	Human Behavior	3
Gen Ed	Option D Course 1	3
Junior Year – Second Semester		
MSIM 331	Simulation Software Design	3
MSIM 383	M&S Laboratory 3	1
MSIM 351	Analysis for M&S	3
MSIM 310	System Modeling	3
Gen Ed	Interpreting the Past	3
Elect	Approved Elective 1	3
Senior Year – First Semester		
MSIM 441	Computer Graphics & Visualization	3
MSIM 487	Capstone Design I	4
Gen Ed	Option D Course 2	3
ENMA 401	Project Management	3
MSIM 4xx	Approved MSIM Elective 1	3
Senior Year – Second Semester		
ENMA 480	Engineering Ethics	3
MSIM 488	Capstone Design II	3
Elect	Approved Elective 2	3
MSIM 4yy	Approved MSIM Elective 2	3
ENGN 401	FE Exam Review	1

Figure 2. Showcase M&SE Curriculum

simulations are allowing scientists and engineers to study increasingly complex systems. However, present-day scientists and engineers frequently lack the basic knowledge of how to effectively apply modeling and simulation within their disciplines. Often existing computer tools are not capable of studying today's systems because the tools were designed only for older system constructs or the tools cannot scale. Therefore, students should have an understanding of basic M&S techniques to allow them to go beyond current tool capabilities. The minor for the user population then focuses on giving students the ability to develop models and simulations using state of the art techniques in both the discrete event and continuous domains. They also learn the proper analysis techniques to run simulations in an experimental environment. Courses available to this population include Discrete Event Simulation, Continuous Simulation, Simulation Analysis, and Systems Modeling.

The M&S developer population is comprised of computer scientists and computer engineers looking to extend their education into the M&S field. This population is characterized by having a strong background in calculus and object oriented programming. In the minor, they then are exposed in some depth to the software development aspects of M&S. Courses available to this constituency include Discrete Event Simulation, Simulation Software Design, Computer Graphics and Visualization, Distributed Simulation, Simulation Analysis, and Game Development.

As basic M&S concepts become more essential and better integrated in the engineering and science disciplines, it is anticipated that the need for a minor in M&S will grow. A minor in M&S gives other academic programs the chance to have their major students exposed to M&S without taking away from the major course content in their present curriculum requirements.

5. GRADUATE PROGRAM IN M&S

The M&S Graduate Program is administered through the MSVE Department and is directed at the M&S developer student constituency. The MSVE Department offers programs of study leading to the degrees Master of Engineering (ME) in M&S, Master of Science (MS) in M&S, Doctor of Engineering (DEng) in M&S, and Doctor of Philosophy (PhD) in M&S. The ME and DEng programs are directed primarily at part-time students employed full-time in the M&S industry who are seeking a more solid foundation in the discipline and/or preparing for technical leadership positions. The MS and PhD programs are directed primarily at full-time students who are preparing for a career in advanced M&S research and/or academic positions.

The ME program is available only as a non-thesis option and is designed around a strong set of core courses addressing the foundation of the M&S Body of Knowledge. Required core courses include: Introduction to M&S; Discrete Event Simulation; Continuous Simulation; Engineering Systems Models; Analysis for M&S; Visualization for M&S; and Distributed Simulation. Three elective courses also are required and allow the student to investigate advanced M&S fundamentals or applications of M&S in various domains. This program is available live, synchronously through two-way television; and asynchronously through the Internet.

The DEng program is available only to M&S industry practitioners having at least two years of engineering experience. Candidates must have the cooperation and support of their employer. The program consists of a core of engineering management and leadership courses and advanced technical M&S courses. An applied project that demonstrates the candidate's ability to apply technical and managerial skills to the solution of a significant engineering problem also must be completed. At present, this program is offered only in the live and televised formats, but plans are underway to offer the program asynchronously via the Internet. Project activity requires periodic live interactions between the candidate and participating faculty and industry supervisors.

The MS program is available only as a thesis option. The curriculum is designed around a reduced set of core courses that includes five courses: Introduction to M&S; Discrete Event Simulation; Engineering System Modeling; Analysis for M&S; and Visualization for M&S. In addition, students must complete a thesis research project and three elective courses designed to support the thesis research. The thesis research is designed to provide a research apprenticeship in which the candidate conducts guided research in an area of M&S. While some of the course work is available via televised instruction and asynchronous web delivery, students are expected to be present on campus to work with their supervising faculty during the completion of their thesis research.

The PhD program focuses on developing the necessary skills and advanced knowledge to evaluate and conduct independent original research in an area of M&S. The goal of the program is to prepare students for careers in teaching and research in academic institutions, as well as the conduct or leadership of research and development in public and private organizations. The program requires the completion of four core courses and four elective courses selected to aid the dissertation research. The program also requires the successful completion of a progressive sequence of program examinations including the Diagnostic Exam, the Qualifying Exam, the Dissertation Proposal, and the Dissertation Defense. Once again, some of the course work is available via televised instruction and asynchronous web delivery, but students are expected to be present on campus to work with their supervising faculty during phases of the dissertation research.

At the present time, there are no undergraduate programs in M&S. Students come to the M&S graduate program having only basic skill sets in mathematics, computer science, and possibly engineering. Therefore, the MS and ME programs are designed for students with little or no background in M&S. As more undergraduate M&S programs are established, some portion of entering M&S students will have significant backgrounds in M&S fundamentals. This will create a new challenge, and an exciting opportunity, for M&S graduate programs. Master's programs will need to be expanded and elevated so that they challenge the students having undergraduate M&S degrees. At the same time, background courses will be needed to allow students entering the program without an undergraduate M&S degree to be successful.

6. CONCLUSION

A new multi-faceted model for university-level M&S academic programs is described. The organizational model simultaneously supports the needs of both the user student constituency and the developer student constituency. An M&S graduate certificate program provides M&S users with the skills and knowledge that they need to apply M&S in the study of other disciplines. An academic department offering a full spectrum of M&S academic degree programs provides M&S developers the opportunity to achieve traditional academic credentials in M&S. The research activities of both student constituencies are further enhanced by the presence of a university-wide M&S research center.

Perhaps the single most important aspect of this model has been the development of an undergraduate program in M&SE. When fully implemented, this program will produce a source of entrance-level engineers that are heavily in demand by the M&S industry. Industry no longer will need to retrain graduates from other academic disciplines to support their M&S activities. The existence of the M&SE bachelor's program also has served as a catalyst in the development of M&S educational programs at the community college level and the high school level. This year, for the first time ever, our state has a workforce development educational pathway in M&S that extends from high school through the doctoral degree. Students interested in building a career in M&S can enter and exit this educational pathway several times as they progress up the technical position hierarchy in M&S.

A second very positive outcome achieved by implementing this organizational model has been the positive response from the M&S user constituency. Students from across the university are discovering and applying M&S. Their presence in graduate classes and their participation on multi-disciplinary research teams has enriched the academic environment. This cross-disciplinary exchange among students now is beginning to encourage faculty to explore the formation of faculty inter-disciplinary and multi-disciplinary research teams to address problems that could not be solved through other approaches. Our experiences so far have been very promising.

REFERENCES

[1] "Simulation-Based Engineering Science", Report of the National Science Foundation Blue Ribbon Panel on Simulation-Based Engineering Science, National Science Foundation, May 2006.

[2] Belytschko, T. and Lazowska, E. (chairs), "Computational Science: Ensuring America's Competitiveness", President's Information Technology Advisory Committee (PITAC) Report, June 2005, web source: **www.nitrd.gov.**

[3] Rogers, R., "What Makes a Modeling and Simulation Professional: The Consensus View from One Workshop," Proceedings of the Winter Simulation Conference, Piscataway, NJ, 1997.

[4] Oren, T.I., "Toward the Body of Knowledge of Modeling and Simulation", Proceedings of the Interservice/Industry Training, Simulation and Education Conference (I/ITSEC), Paper 2025, Orlando, FL, December 2005.

[5] Petty, M.D., "Graduate Modeling and Simulation Overview Course," Final Technical Report, Contract N00014-031-0948, Defense Modeling and Simulation Office, March 2006.

[6] Sarhoughian, H.S. and Zeigler, B.P., "Towards Making Modeling and Simulation into a Discipline,' 2000, web source: **www.acims.arizona.edu.**

[7] Nance, R.E., "Simulation Education: Past Reflections and Future Directions", Proceedings of the 2000 Winter Simulation Conference, pp. 1595-1601, 2000.

[8] Nance, R.E. and Balci, O., "Thoughts and Musings on Simulation Education", Proceedings of the Winter Simulation Conference, pp. 1567-1570, 2001.

[9] Roberts, C.A. and Ghosh, S., "A Proposed Model for an Undergraduate Engineering Program in Modeling and Simulation", Proceedings of the International Conference on Simulation Education, Western Simulation Multi-conference, pp. 38-44, 2004.

[10] Szczerbicka, H., Banks, J., Rogers, R., Oren, T., & Sarjoughian, H.Z., "Conceptions of Curriculum for Simulation Education," 2000 Winter Simulation Conference, pp. 1635-1644, 2000.

[11] Sarhoughian, H., Cochran, J., Collofello, J., Goss, J., and Zeigler, B. "Graduate Education in Modeling & Simulation: Rationale and Organization of an Online Masters Program," Proceedings of the Summer Computer Simulation Conference, 2004.

[12] Mielke, R., Scerbo, M., Gaubatz, K., and Watson, G., "A Model for Multidisciplinary Graduate Education in Modeling and Simulation,". International Journal of Simulation and Process Modeling, Vol. 5, No. 1, pp. 3-13, February 2009.

[13] University of Alabama – Huntsville, web source: **http://catalog.uah.edu,** accessed May 25, 2011.

[14] Arizona State University, web source: **http://www.asu.engineeringonline.com,** accessed May 25, 2011.

[15] California State University – Chico, web source: **http://www.ecst.csuchico.edu/~mcleod/degree/html**, accessed May 25, 2011.

[16] Georgia Institute of Technology, web source: **http://www.msrec.gatech.edu,** accessed May 25, 2011.

[17] Old Dominion University, web source: **http://eng.odu.edu/msve**, accessed May 25, 2011.

[18] University of Central Florida, web source: **http://www.ist.ucf.edu/phd**, accessed May 25, 2011.

[19] Petty, M., "Modeling and Simulation Body of Knowledge Topics", in Certified Modeling & Simulation Professional Program – Management Plan, Modeling and Simulation Professional Certification Commission, Version 1.0, June 12, 2009, from **http://www.simprofessional.org.**

[20] James F. Leathrum, Jr. and Roland R. Mielke, "A Bachelor of Science in Modeling and Simulation Engineering", International Symposium on Engineering Education and Education Technologies, Orlando, FL, July 19-22, 2011.

[21] Birta, L.G., "The Quest for the Modelling and Simulation Body of Knowledge," Proceedings of the Sixth Conference on Computer Simulation and Industry Applications, Tijuana, Mexico, 2003.

[22] ABET Engineering Accreditation Commission, "Criteria for Accrediting Engineering Programs", 2011-2012 Accreditation Cycle, web source: **http://www.abet.org,** accessed May 23, 2011.

Return on Investment: A Practical Review of Learner Modeling Techniques

AUTHORS

Jeremiah T. Folsom-Kovarik
Institute for Simulation and Training, University of Central Florida, Orlando, Florida, USA
(321) 947–7246
jfolsomk@ist.ucf.edu

Sae Schatz
Institute for Simulation and Training, University of Central Florida, Orlando, Florida, USA
(407) 920-4981
sschatz@ist.ucf.edu

ABSTRACT

Adaptive instructional systems, such as intelligent tutors, personalize instruction for individual learners. As such, their learner models are critical to their performance; yet, creation of these models can contribute significantly to development costs. This article reviews published findings about learner models from the perspective of balancing potential learning effects against creation costs. Other issues considered include models' abstraction, fitness for open-ended or ill-defined domains, generalization to new problems, and macro- versus microadaptation.

KEYWORDS

Intelligent tutoring, learner models, return on investment, development costs, business case

ACKNOWLEDGEMENT

This work is supported in part by the Office of Naval Research Grant N0001408C0186, the Next-generation Expeditionary Warfare Intelligent Training (NEW-IT) program. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the ONR or the US Government. The US Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

ABOUT THE AUTHORS

MR. JEREMIAH T. FOLSOM-KOVARIK is a Research Assistant with the UCF-IST Applied Cognition and Training in Immersive Virtual Environments (ACTIVE) lab. He is working toward a computer science Ph.D., and his research focuses on novel interaction and modeling opportunities in intelligent tutoring systems.

DR. SAE SCHATZ is now a Human–Systems Scientist at MESH Solutions, LLC, a DSCI Company. Dr. Schatz performs applied research in cognitive readiness, military instruction, scenario-based training, and adaptive instruction.

INTRODUCTION

Personalized instruction can be highly effective. For instance, students who learn from a dedicated human tutor perform substantially better than those who are educated in group-oriented classroom settings (Bloom, 1984). However, one-on-one tutoring is expensive—requiring many more resources per student than teaching a group. One way to help personalize learning, without requiring more instructors, is to supplement human attention with adaptive computerized tools.

Adaptive instructional technologies, such as Intelligent Tutoring Systems (ITSs), personalize instruction for each student. ITSs can respond to students' needs by detecting mistakes and correcting them, customizing their teaching styles, or even changing what material they teach. To make this adaptation possible, most ITSs contain internal models that reflect what they "know" about the student, the material, and how to teach. In particular, the learner model (also called the student, trainee, or user model) lets an ITS compile information about learners and respond to them. As such, learner models are critical components of adaptive instructional systems.

In an ideal world, learner models would only be designed to best support instructional objectives, but in reality, development costs must also be considered. Creation of learner models can represent up to a third of ITSs' planning and development costs (Folsom-Kovarik,

Schatz, & Nicholson, 2010). Developing a sophisticated learner model, for instance, may require the equivalent of five people working full-time for a year (Koedinger, Aleven, Heffernan, McLaren, & Hockenberry, 2004), while simpler models may be ready in just a few days (Blessing, Gilbert, Ourada, & Ritter, 2009; Folsom-Kovarik et al., 2010). Thus, when designing a new adaptive instructional system, there is a clear need to balance models' costs and benefits. To help practitioners do this, we have conducted a review of learner modeling approaches and attempted to identify measures of both their effectiveness and costs; we have also summarized three key considerations that further affect the type of model to be used. Although multifaceted concepts (like learner model cost-benefit ratios) cannot be reduced to simple universal answers, we hope this review offers meaningful insights into learner modeling from a practical perspective.

INTELLIGENT TUTORING SYSTEMS (ITSS)

Different forms of adaptive instructional systems exist, and broadly speaking, they all attempt to give different learners different help. In general, the premise of adaptive instructional systems is that personalized instruction improves learning effectiveness and, often, instructional efficiency. Intelligent tutors are among the most well-studied adaptive instructional technologies. They employ artificial intelligence and an array of modeling techniques to adapt instructional content and its delivery to each student's needs—similar to how human tutors adaptively interact with their students but without the need for direct instructor involvement.

ITSs have been used in a variety of real-world settings, from avionics maintenance to grade-school mathematics. Simple ITSs can be presented via the web, while more complex versions employ sophisticated artificial intelligence and require dedicated software platforms. More recently, intelligent tutoring mechanisms have also been integrated with simulation-based training environments (for a review see Schatz, Oakes, Folsom-Kovarik, & Dolletski-Lazar, in press).

ITSs' effectiveness depends on their underlying models. Heuristically, most ITSs include three models: (1) *a domain model* that organizes and articulates the subject matter (that is, "what to teach"), (2) *a pedagogical model* that describes instructional approaches (or "how to teach"), and (3) *a learner model* that describes the unique attributes and performance history of students so that the software can determine learners' strengths and weaknesses, then select content and teaching approaches that meet each learner's unique needs. The remainder of this paper focuses on learner models and the variety of modeling techniques that can implement them.

SURVEY OF LEARNER MODELS

This section describes a variety of learner model types, from the most detailed to the most abstract. When available, quantitative information about models' learning effects and typical development costs are provided. For comparability, learning outcome data are reported as *effect sizes*, which describe the difference in pretest–posttest improvement between a control group and an experimental group. For example, an effect size of 1.0 means that the average improvement in the experimental group was one standard deviation higher than the average improvement in the control group (Schulze, 2004). Figure 1 illustrates an example effect size of 0.8; in this sample graph, the experimental group's average score is almost a full grade-level higher than the control group's average. As a reference point, consider that human tutors can produce an effect size of about 2.0 as compared with large-group instruction (Bloom, 1984).

The development costs of the systems are reported as a ratio of person-hours to hours of instruction. The ratio format makes cost figures more comparable across different ITSs that may undertake more or less complex tutoring tasks.

Production rule models

The most detailed models currently used in ITSs are the *production rule* systems that employ a *model tracing* algorithm. These ITSs contain many rules drawn from a general model of human cognition. The ITS traces learner cognition by finding the rules that produce the same outcomes as the learner did. For example, if there is a 30-second delay after an ITS asks a question before the learner responds to it, a model-tracing tutor might contain rules that estimate how much of that time the learner spent reading the question, how much selecting an answer, and whether the delay indicated confusion, mental load, and so on. Model-tracing tutors attempt to capture students' cognitive steps at a very granular level, so they are often called *cognitive tutors* (Anderson, 1993). Many cognitive tutors use the ACT, ACT*, or ACT R cognitive model, although some

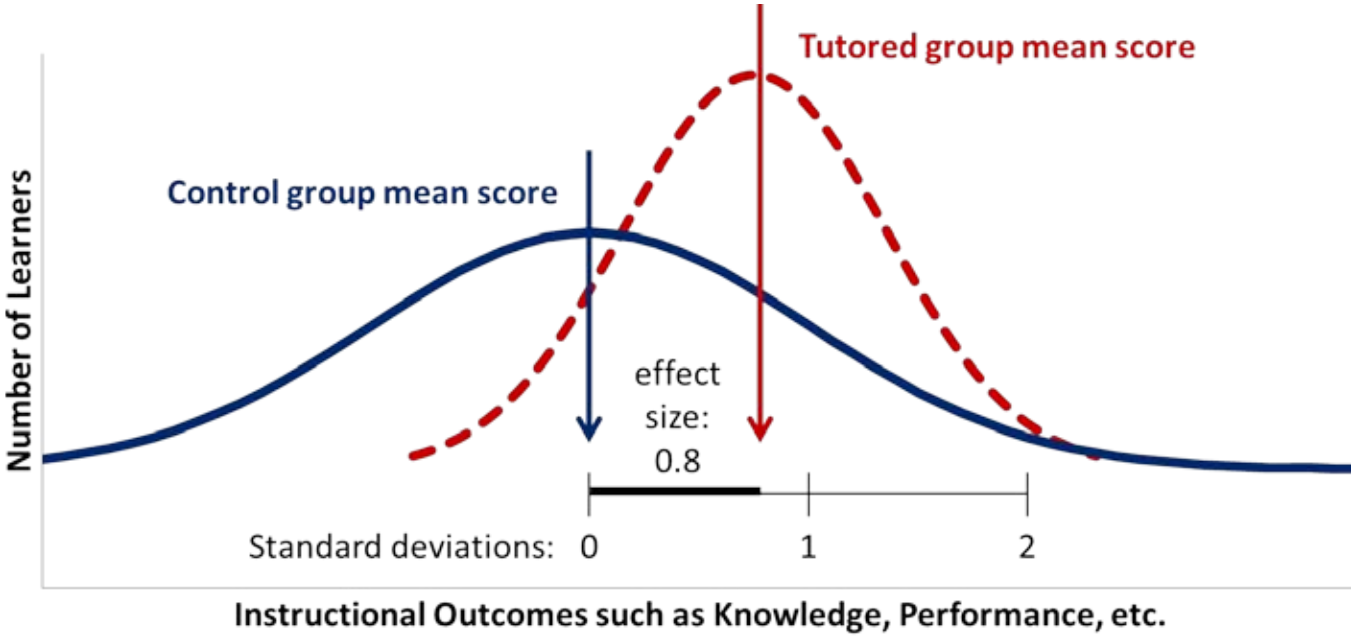


Figure 1. *Effect size* describes the difference between two bell curves.

production-rule systems have been based on other cognitive theories (Callaway et al., 2007; B. G. Johnson, Phillips, & Chase, 2009; Neches, Langley, & Klahr, 1987).

Early model-tracing systems required between 100 and 1000 hours of development for each hour of instruction, which was comparable to the time required to build other computer-aided instruction systems of the time (Anderson, 1993, p. 254). More recently, developers of a model-tracing algebra tutor reported that its development took approximately 10,000 hours and provided 50 hours of instruction, or 200 development hours for each hour of instruction (Koedinger et al., 2004). Simple model-tracing tutors can also be built with the Cognitive Tutor Authoring Tools (CTAT), a set of tools designed to speed the authoring process. A formative study, where four graduate students built a model with just six rules, showed that using CTAT sped up the development process by 40 percent (Aleven, McLaren, Sewall, & Koedinger, 2006). If this result can generalize to larger projects, future model-tracing tutors might require less than 100 hours of development time for each hour of instruction.

Early model-tracing tutors demonstrated success in several settings. A high-school geometry tutor produced improvements of "more than one standard deviation" compared to classroom study (Anderson, Corbett, Koedinger, & Pelletier, 1995, p. 183), and a tutor for teaching college students the LISP programming language produced an effect size of 0.75 compared to working the problems without hints (Corbett, 2001; Corbett & Anderson, 2001). Finally, in conjunction with an overhauled curriculum, the PUMP Algebra Tutor (PAT) taught algebra skills to a large sample of ninth-grade students in urban schools. Compared to students who received all their math instruction in a classroom, students who used PAT during some class periods displayed improved learning effect sizes of 0.7 to 1.2 (Koedinger, Anderson, Hadley, & Mark, 1997).

As shown in this section, production-rule tutors have been fielded and tested in meaningful real-world settings, and they generally engender substantial learning effects. In fact, production-rule tutors have among the highest rates of effectiveness of all the tutor varieties discussed in this paper. However, even the most economical production-rule tutors require a substantial development time investment. Some researchers who use model-tracing systems argue that implementing a specific cognitive theory is worth the added cost (Neches et al., 1987), and empirical results generally show that cognitive tutors can produce learning increases among the best of any ITS.

General perturbation models

A broad class of learner models, called *perturbation models* or *buggy models*, try to describe all the incorrect

knowledge the learner may have. Incorrect knowledge, variously called misconceptions, mal-rules, or bugs, represents persistent errors in thinking that the ITS should correct. The model-tracing ITSs described above could be considered relatives of buggy models because their production rules also include rules for incorrect actions and misconceptions (VanLehn, 1988). In general, buggy models do not necessarily use any particular cognitive model or method to interpret evidence of bugs. As an example, a buggy-model ITS for teaching algebra might contain a library of common algebra mistakes that let it recognize why learners give wrong answers on an algebra test.

While some buggy models can be automatically populated based upon cognitive theory (e.g., DEBUGGY, (Brown & VanLehn, 1980); IDEBUGGY, (Burton, 1982)), other systems require subject matter experts to manually enumerate an exhaustive list of bugs (e.g., PROUST, (W. L. Johnson, 1990)). Therefore, buggy models can entail a high cost, and this cost may grow considerably if developers hope to generalize the model to different learner populations (Payne & Squibb, 1990). To provide a specific figure, one report on buggy-model development gave a cost ratio of 133:1 (Folsom-Kovarik et al., 2010), which is in line with production rule system costs.

In terms of performance, published results suggest that buggy models have performed somewhat less well than production-rule models. One buggy-model ITS, called Smithtown, let students explore and manipulate an artificial economy (Shute & Glaser, 1990). Students who spent five hours with the ITS did as well on a test of economic principles as a control group of students who spent eleven hours studying in a classroom. More recently, a learning system called Adaptive Content with Evidence-based Diagnosis (ACED) used a buggy model for error feedback while teaching algebra to high school students (Shute, Hansen, & Almond, 2008). ACED did show an effect size of 0.38, but only in comparison to no intervention at all; although it should be noted that the main focus of the ACED project was not on improvement in learning but on improving assessment without damaging students' learning.

Finally, two dialog-based tutors, AutoTutor (Graesser et al., 2004; Graesser, Wiemer-Hastings, Wiemer-Hastings, & Kreuz, 1999) and Atlas (VanLehn et al., 2002), both estimate students' knowledge with short lists of common misconceptions—on average, fewer than five (VanLehn et al., 2007). These tutors have achieved effect sizes as high as 1.02 (Graesser et al., 2003; VanLehn et al., 2007). However, these learning outcomes are usually attributed to the unique dialogue interaction of the tutors, rather than their very simple learner models (Rosé et al., 2001).

In sum, detailed buggy models require significant effort to build, and typical models have not been developed at low costs. Furthermore, the published literature lacks examples of buggy-model ITSs that demonstrated a substantive improvement in learning effectiveness, beyond the dialogue-based systems mentioned above. For these reasons, buggy models (when used as the sole learner model in traditional ITSs) may not offer sound return on investment.

Example tracing

Like tutors with buggy models, *example-tracing* tutors can respond to errors that learners make at a detailed level. However, example-tracing tutors do not model misconceptions in a list that can generalize to multiple situations. Instead, system authors define incorrect responses for single questions, and they are less concerned with the cognitive theories that underlie these mistakes. An example-tracing tutor might ask a multiple-choice question and have responses built in for each of the wrong answers, but not have any programming to deal with other, similar questions. This abstraction away from a detailed cognitive theory also gave example-tracing systems their old name, pseudo-intelligent tutors or *pseudotutors*.

Example tracing models were created in direct response to the high development costs of the model-tracing approach (Koedinger et al., 2004). From their inception, development of example-tracing systems has been sped by the CTAT authoring tools, which are also used to create model-tracing systems. In preliminary tests, domain experts needed an average of about 23 hours of design and development time to create one hour of example-tracing instruction with CTAT (Koedinger et al., 2004), and in recent studies, novices required only 30 (Heffernan et al., 2006) to 40 hours of development time to create one hour of instruction (Razzaq et al., 2008).

Compared to pen-and-paper homework, the ASSISTment example-tracing tutor produced a learning improvement of 0.61 standard deviations (Mendicino,

Razzaq, & Heffernan, 2009). Another example-tracing tutor, which taught logic puzzles to college students, produced an effect size of about 0.75 compared to pen-and-paper practice (Hockenberry, 2005), and this system (built with CTAT) required only 18 hours of development for one hour of tutoring time.

Example-tracing models are less detailed and less general than canonical cognitive tutors and buggy-model tutors. This may explain why even the best example-tracing tutors have only produced moderate learning gains, roughly 60% of the effectiveness reported for full production-rule systems. However, published experiences show it is possible to rapidly build an example-tracing tutor, and for this reason they offer a high cost-benefit ratio.

Constraint-based modeling

Constraint-based modeling eschews extensive models of learner cognition and instead constructs libraries of domain-relevant constraints against which learners' actions are compared (Ohlsson, 1994). Constraint-based models need not track historic performance or even specific user actions, but instead monitor the immediate problem state. As long as a learner never reaches a state that the model identifies as wrong, he or she may perform any action. This allows constraint-based tutors to selectively abstract away certain details. As an example, an ITS using constraints in a simulation environment might let learners wander around an area for some time without trying to model their specific actions, but would intervene when they take too long, move into a restricted area, or otherwise make an outright mistake.

An early constraint-based tutor that taught SQL to graduate students had a cost ratio of 220:1 (Mitrovic & Ohlsson, 1999), but since that time, several authoring tools have been created that improve the efficiency of their development. For example, using the Constraint Authoring System (CAS), graduate student participants were able to build a constraint-based tutor for adding fractions in, on average, 31.3 hours (Suraweera, Mitrovic, & Martin, 2007). The Web-Enabled Tutor Authoring System (WETAS) similarly helped four graduate students work together to create a small spelling tutor, requiring 32 person-hours to complete (Martin, Mitrovic, & Suraweera, 2008). Unfortunately, the teaching times of these small tutors were not reported.

Several well-studied ITSs have used constraint-based methods, giving a good overview of the typical learning effects they can achieve. For instance, a constraint-based tutor that taught SQL to college students produced an effect size of 0.63 as compared to an ablated version that used no adaptation (Suraweera & Mitrovic, 2004). Another constraint-based tutor in the same domain had previously produced an effect size of 0.75, but in that formative study selection bias may have confounded the results (Mitrovic & Ohlsson, 1999). A third constraint-based ITS taught a small group of students collaboration skills. When the ITS provided immediate feedback, knowledge about collaboration tactics improved by 1.3 standard deviations, compared to students using the ITS with no collaboration feedback (Baghaei & Mitrovic, 2007).

Finally, in an interesting case, the Andes physics tutor was first created with production rules and a Bayesian network (described below), but it was later remade with constraints (VanLehn et al., 2005). Evaluated annually over four years, Andes with a constraint-based learner model yielded an overall effect size of 0.61 compared to working practice problems on pen and paper. The learning gain was especially large, 0.70 and 1.21, in the two areas where the Andes material most closely aligned with the course test material (VanLehn et al., 2005). This effect size is very close to the result reported for the model-tracing PAT tutor described above.

Overall, these development costs and outcomes are similar to those of production-rule systems—the most consistently effective, but also the most expensive, models. However, unlike production-rule models, simpler models could be created with the constraint-approach, but the development costs and learning outcomes of such simple constraint-based systems have not yet been recorded.

Bayesian networks and other classifiers

A Bayesian network consists of a collection of known values, such as how well a learner is performing on a particular test question, and estimated values, such as how well that learner understands the underlying concepts. The model includes relationships between known and unknown values that make its predictions reasonable (Charniak, 1991). For example, a Bayesian model could take a learner's wrong answer to one question and infer that the learner is more likely to make a mistake on several other questions. Like the other classifiers described in this section, Bayesian networks often model few details about

learners, because Bayesian networks become more difficult to adjust as their size increases. A Bayesian network large enough to differentiate the hundreds of detailed misconceptions that some buggy libraries use would be difficult to initialize, and its estimates would become highly suspect (Ott, Imoto, & Miyano, 2004).

Wayang Tutor used a Bayesian network to interpret data, such as the correctness of answers and hint selection, without tracing every step of problem solving (Arroyo, Woolf, & Beal, 2006). One study showed that the tutor helped students learn, but the control and experimental groups were dissimilar and no effect sizes could be given (Beal, Walles, Arroyo, & Woolf, 2007). Another study found Wayang Tutor was as effective as small-group study with a human tutor (Beal, Shaw, & Birch, 2007), and comparing Wayang Tutor to classroom instruction or with non-interactive websites yielded an effect size of 0.39 (Arroyo, Woolf, Royer, Tai, & English, 2010).

Dynamic Bayesian networks are Bayesian networks that can account for change over time. The Prime Climb educational math game used a dynamic Bayesian network to model students' affect and then adapt its hints accordingly. The game showed a learning effect size of 0.7 as compared to students who played the same game with no hints at all, and a "modified roll-up" method made the design of the dynamic network only slightly more complicated than the design of a static one (Conati & Zhou, 2004).

Other static and machine-learning classifiers can also play the role of a student model. Examples of such classifiers that have been used as student models include *finite-state automata* (e.g., Stottler, Fu, Ramachandran, & Vinkavich, 2001), *decision trees* (e.g., Cha et al., 2006; McQuiggan, Mott, & Lester, 2008), *neural networks* (e.g., Castellano, Mastronardi, Di Giuseppe, & Dicensi, 2007), *case libraries* (e.g., Kass, Burke, Blevis, & Williamson, 1994; Reyes & Sison, 2002), and ensemble methods (e.g., Hatzilygeroudis & Prentzas, 2004; Lee, 2007). In at least some practical situations, many classifiers are interchangeable in their performance (McQuiggan et al., 2008; Walonoski & Heffernan, 2006).

Published reports typically do not detail the development effort needed to create classifiers. In part, this may reflect researchers' perception that classifiers are "off-the-shelf" technology and their implementation can be completed with little effort. As an example, several military tactics tutors created with the Internet ITS Authoring Tool needed "a small fraction of the time normally required" to create learner models (Stottler et al., 2001, p. 1). Similarly, the Cognitive Model SDK is an authoring tool for manually developing hierarchical rules whose predicate sets function similarly to decision trees. This architecture let undergraduate novices develop the model for a fraction addition tutor in 7.68 hours on average, or about a quarter of the time novices in a separate study needed to develop a constraint-based model in CAS for a similar tutor (Blessing et al., 2009). Finally, one anonymous practitioner reported building a classifier model which required 30 hours of develop per hour of instruction, while another reported spending 50 hours per instruction hour (Folsom-Kovarik et al., 2010).

In summary, classifiers collectively represent an efficient approach to developing learner models. However, they typically have lower detail and have produced lower learning effectiveness outcomes than some of the more detailed systems, such as that use the production-rule, constraint-based, and example-tracing approaches. Nonetheless, due to their low development costs, classifier systems have good cost-benefit ratios.

Overlay models

Many ITSs, especially early examples such as Scholar (Carbonell, 1970), PLATO West (Burton & Brown, 1976), and Wusor II (Carr, 1977, p. 66), modeled learners' knowledge with an *overlay*. Overlay models ignore details of *how* students learn and instead track *what* students have learned in a simple way, similar to a checklist. An example would be an ITS that re-teaches several target skills until a learner demonstrates each skill once. Overlays' high degree of abstraction, however, does not lead to flexibility in learner interactions. On the contrary, overlays tend to force learners into specific answers and discount learner knowledge that falls outside the ITS's model of expert knowledge (Burton & Brown, 1976).

Developing overlay models requires expert knowledge of the domain in order to specify topic definitions, prerequisites, and ordering. After knowledge elicitation about the content, though, there are few technical challenges to building overlay models. Practitioners have been able to spend as little as 24 development hours per instruction hour, although two other projects reported ratios of 100:1 or 667:1 (Folsom-Kovarik et al., 2010). These figures remind us that while some model types can be developed

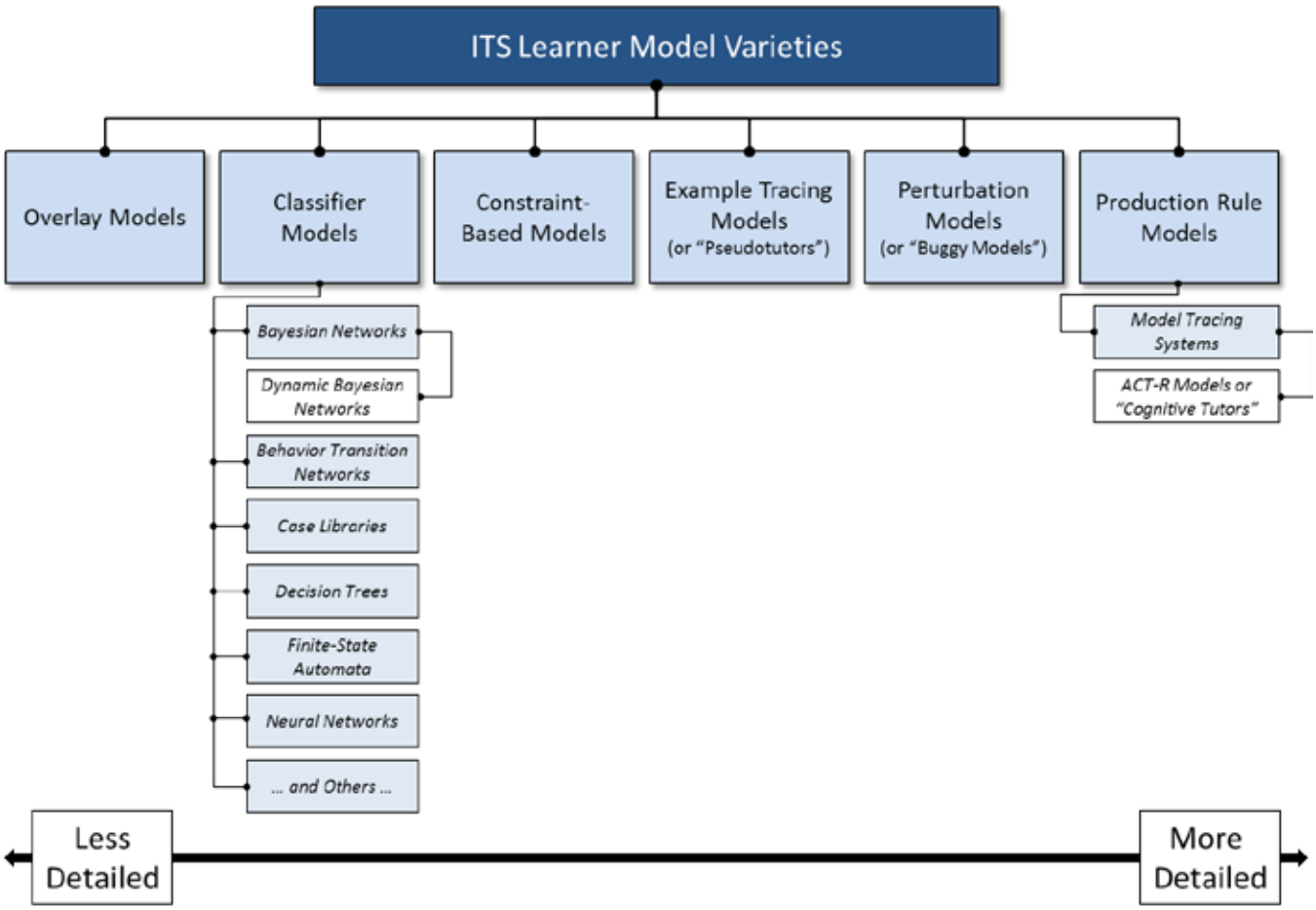


Figure 2. While the relationships between models are complex, a notional grouping helps contextualize the types discussed.

at low cost, any modes can have high costs under different circumstances.

One example of an overlay model is the classic ITS Sherlock, which taught electronics troubleshooting. Sherlock implemented a rather intricate overlay model that could estimate trainees' abilities as one of four levels for each skill (Lesgold, Lajoie, Bunzo, & Eggan, 1988). The model also allowed for learning and forgetting. Comparing trainees who worked for 20 hours with Sherlock against those who had 20 hours of on-the-job instruction gave an effect size of 1.02 (Shute, 1990).

In summary, although the models' development costs can vary widely, it is possible to develop overlay models quite inexpensively. Overlay models have produced moderate effect on learning in the past. The effectiveness and development costs of overlay models are likely to vary based on the effort and subject matter expertise dedicated to each instance. Further, few modern systems use pure overlays. This is because, as the following section discusses, several modern design considerations often limit overlay models' instructional effectiveness.

OTHER CONSIDERATIONS

The previous section briefly surveyed six common categories of ITS learner models. Their degrees of abstraction, learning impacts, and development costs were described (see Figure 2). However, selection of a learner model type will be affected by other important considerations, too. These include whether the subject matter is *poorly defined* or unusually *open-ended*, how much *generalizability* is required, and whether *macroadaptation* or simply *microadaptation* will be employed.

Open-ended and ill-defined domains

How well defined a domain is affects the clarity of the rules the tutor can use to evaluate students' actions. Some domains are difficult to distill into clear-cut rules.

For example, social interaction tasks call on subtle experiences that may be difficult to program into a computer (Kass et al., 1994). Relatedly, domains in which a learner can make many diverse decisions, or in which the same decision can lead to many distinct actions, may be difficult to operationalize.

Two learner model types work better in open-ended or ill-defined domains because they place fewer limits on the types of actions learners can take. First, in constraint-based models, learners may perform any action as long as they never reach a state that is constrained in the model. Because of this permissiveness, constraint-based approaches may offer the best solution for open-ended problem domains (Mitrovic & Weerasinghe, 2009). Second, case libraries, a subset of the classifier models, have an advantage in domains that are too complex, subjective, underexplored, or otherwise imprecisely defined to specify exactly (Lynch, Ashley, Aleven, & Pinkwart, 2006).

Generalization

A model's ability to generalize relates to its software modularity and reusability. More generalizable models allow greater reuse, which may involve adding new material or repurposing the ITS for a new audience. Models may lose generalizability if they contain built-in assumptions about the material, students, or pedagogy. Generalization and reusability are weaknesses of most contemporary ITSs. In their ideal implementations, production-rule and constraint-based models should allow for some reuse, although typical implementations of both sacrifice generalization for development expediency (Kodaganallur, Weitz, & Rosenthal, 2006).

The most basic way to generalize or repurpose an ITS is to add more problems similar to those already in the tutor. This task is simplest for buggy, model-tracing, and constraint-based tutors. Their models typically use generic rules or constraints that still apply to new problems, with few or no changes needed (Kodaganallur et al., 2006). Overlays, Bayesian networks, and other classifiers may need small changes to address the new material. Finally, example-tracing systems are the least able to handle new content: even when new problems have the same structure and meaning as existing items, they still require new examples to describe correct and buggy solutions.

Another reuse task that practitioners might face is redeploying an ITS to a new audience. The misconceptions included in buggy models can vary significantly among populations, even between one school and another (Payne & Squibb, 1990; VanLehn, 1982). Moving between groups with larger differences may cause performance degradation, and since model-tracing systems are related to buggy models, their reusability might also be questioned. Some model-tracing rules may need to be completely rewritten for new populations (Kodaganallur et al., 2006; Mitrovic, Koedinger, & Martin, 2003), although the widespread use of some model-tracing systems suggests that their production rules are actually fairly reusable (Koedinger et al., 1997). Constraint-based tutors may be more resilient still (Mitrovic et al., 2003); however, in practice, some constraint-based tutors do use constraints equivalent to production rules, so there may be little practical difference in the ease of redeploying constraint-based and production-rule systems (Kodaganallur et al., 2006). Finally, systems that incorporate machine learning, include some case-based, Bayesian, and other classifier models, are all designed to adapt to new student populations within the parameters of their machine learning algorithms (Murray, 2003). Thus, these systems should boast good capacities for generalization and reuse with little additional effort.

Macro- and microadaptation

Although all ITSs adapt, or personalize, their interactions with students, the timing of their adaptations can differ. Snow and Swanson (1992) identify three options for when to diagnose and subsequently adapt the system. First, these can occur prior to a training episode based upon pre-task measures or historical data; this is called *macroadaptation*. Second, adaptive interactions can take place during the training task based upon ongoing performance or behavioral assessment; this is called *microadaptation*. Finally, an ITS can use macro- and microadaptation in combination.

In modern parlance, systems that use macroadaptation alone are likely to be termed computer-based training systems or adaptive learning systems, while many intelligent tutors use microadaptation only (Ong & Ramachandran, 2005). In some environments, there may be practical reasons to avoid implementing macroadaptation. For example the Andes physics tutor, which in its original version used a Bayesian network to adapt lesson plans (Conati, Gertner, & VanLehn, 2002), was redesigned without macroadaptation because the educators using the system required a specific curriculum presented in a specific order (VanLehn et al., 2005). However, when

feasible, the combination of micro- and macroadaptation will likely have the most impact on learning (Shute, 1992).

Macroadaptation

In the absence of special considerations, a simple overlay model is sufficient to estimate learners' knowledge for macroadaptation. Several other model types described here are more sophisticated, but do not seem to substantially improve the learning effects of a macroadaptive ITS. Further, two model types described here cannot directly support macroadaptation.

First, model-tracing systems do not track the relationships between abstract skills and problem progress (Anderson et al., 1995), so modern model-tracing tutors typically require a secondary model to handle macroadaptation. When used in conjunction with model-tracing, practitioners call this kind of macroadaptive model "knowledge tracing." Knowledge tracing in the ACT Programming Tutor for LISP, a typical implementation, is accomplished with an overlay model that is updated using Bayesian probabilities (Corbett & Anderson, 1995). An early comparison of students using the LISP tutor either with or without the knowledge-tracing component showed assigning problems with knowledge tracing resulted in improved learning with an effect size of 0.89 (Corbett, 2001). Knowledge tracing significantly improves the impact of cognitive tutors, but it also requires additional development work—one practitioner estimated that knowledge tracing took 48 hours to develop for each hour of interaction, while another estimated a 450:1 cost ratio (Folsom-Kovarik et al., 2010). Furthermore, the best way to implement knowledge tracing remains an active research question (Baker, Corbett, & Aleven, 2008).

Like model-tracing systems, constraint-based tutors do not directly model the user, and as such, cannot support traditional macroadaptation. However, practitioners have repurposed constraint sets to act similarly to overlay models, giving a constraint-based tutor the ability to select which problem to present. In one instance, using constraints in this way doubled students' rate of improvement between the first and the second time the tutor tested a constraint (Martin & Mitrovic, 2002).

Microadaptation

To deliver effective microadaptation, a learner model should contain sufficient detail to model learner actions within a single task. More abstract models are less able to support microadaptation, and consequently, overlay models and classifiers are poor choices. Example-tracing systems offer moderate microadaptive capabilities, and they are attractive for their low minimum costs. These may be a good choice for deploying a new ITS quickly or in a series of quick development cycles. However, for long-running or widely deployed projects, example-tracing systems have a disadvantage because they cannot be easily generalized to new contexts. Buggy models, in general, also suffer from generalization issues, and in addition there is little published support for their efficacy.

For involved projects, model-tracing or constraint-based tutors may represent the best microadaptive models. Both of these tutors boast excellent instructional outcomes, and both offer good reusability. Constraint-based tutors may have a slight edge in development costs over model-tracing tutors, and it is possible that they would be better suited to some projects that involve poorly defined or open-ended domains. For these reasons, both model-tracing and constraint-based tutors are good choices for many ITS projects, and constraint-based systems may be the (slightly) better option under certain conditions.

CONCLUSION

An ITS is a teaching tool, designed to enhance students' learning beyond other methods of instruction (such as completing homework or participating in classroom-based lectures). As a collective technology, ITSs have been quite successful, and students who use these adaptive software systems often perform around one letter-grade better than those receiving conventional educational interventions or non-adaptive computer-based instruction.

Intelligent tutors contain learner models, which enable the systems' adaptations. Broadly speaking, the most effective learner model types appear to be constraint-based and production-rule (model-tracing) systems. Overlay models have also been effective in limited contexts such as closed, clearly defined domains where instruction requires mainly macroadaptive interactions.

Intelligent tutors, however, can be costly to create. Highly effective learner models built with constraints and production rules have particularly high development costs. In contrast, example-tracing and classifier-based systems can be developed for relatively little cost, sometimes an order of magnitude less than the more complex

Student model	Lowest reported development to learning time ratio	Highest reported effect on learning	Typical model detail	Open-ended and ill-defined domains	Generalization and reuse	Macro- and micro-adaptation
Production rules and model tracing	200:1	1.2, compared to classroom learning	High	Poor	Good	Micro
Perturbation and buggy models	133:1	Not significant	High	Poor	Some	Both
Example tracing	18:1	0.75, compared to paper homework	Moderate	Poor	Poor	Micro
Constraint-based models	220:1	1.3, compared to work without feedback	Moderate	Good	Good	Micro
Bayesian networks and other classifiers	30:1	0.7, compared to work without feedback	Low	Poor; good with case library approach	Some; good with machine learning approaches	Both
Overlay models	24:1	1.02, compared to on-the-job training	Low	Poor	Poor	Macro

Table 1. Heuristic data on ITS learner model types

constraint-based and production-rule systems. Hence, although they are, perhaps, not the most effective learner modeling approaches, depending on project needs, example-tracing and classifier-based systems may offer the best return on investment.

Table 1 summarizes the characteristics of ITS learner models discussed in this article; this information should be interpreted heuristically. Although all of the data in the table are derived from published accounts of empirical investigations (cited in the text of this paper), we have only reported the best values derived from various sources—the lowest reported development costs and highest reported effect sizes. Additionally, given the number of variables affecting a learning system's success and creation costs, the values presented can only offer rough, "ballpark" insights for future development efforts.

Limitations

Many factors contribute to the learning gain an ITS can realize—from the range of interventions at its command (e.g., Koedinger & Aleven, 2007; Wang, Johnson, Rizzo, Shaw, & Mayer, 2005) to the conduciveness of the learning environment outside the tutor (Kirkpatrick & Kirkpatrick, 2006, pp. 23-25). We make the simplifying assumption that the practitioners creating a new ITS will improve these other factors to the greatest extent possible and only discuss the choice of a learner model for the new system. We assume that the successes of each ITS as a whole can represent a proxy by which to judge its learner model. Of course, in reality the causation is far less direct. Despite these limitations, we hope this succinct review gives valuable real-world information that helps practitioners make decisions about which type of learner model to select when developing a new adaptive instructional system.

REFERENCES

Aleven, V., McLaren, B. M., Sewall, J., & Koedinger, K. R. (2006). *The cognitive tutor authoring tools (CTAT): Preliminary evaluation of efficiency gains*. Paper presented at the 8th International Conference on Intelligent Tutoring Systems.

Anderson, J. R. (1993). *Rules of the Mind.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *Journal of the Learning Sciences,* 4(2), 167–207.

Arroyo, I., Woolf, B. P., & Beal, C. R. (2006). Addressing cognitive differences and gender during problem solving. *Technology, Instruction, Cognition & Learning,* 4, 31–63.

Arroyo, I., Woolf, B. P., Royer, J. M., Tai, M., & English, S. (2010). *Improving Math Learning through Intelligent Tutoring and Basic Skills Training*. Paper presented at the 10th International Conference on Intelligent Tutoring Systems.

Baghaei, N., & Mitrovic, A. (2007). *From Modelling Domain Knowledge to Metacognitive Skills: Extending a Constraint-Based Tutoring System to Support Collaboration*. Paper presented at the 11th International Conference on User Modeling.

Baker, R. S. J. d., Corbett, A. T., & Aleven, V. (2008). *More Accurate Student Modeling Through Contextual Estimation of Slip and Guess Probabilities in Bayesian Knowledge Tracing*. Paper presented at the 9th International Conference on Intelligent Tutoring Systems.

Beal, C. R., Shaw, E., & Birch, M. (2007). *Intelligent tutoring and human tutoring in small groups: An empirical comparison*. Paper presented at the 13th International Conference on Artificial Intelligence in Education.

Beal, C. R., Walles, R., Arroyo, I., & Woolf, B. P. (2007). On-line tutoring for math achievement testing: A controlled evaluation. *Journal of Interactive Online Learning,* 6(1), 43–55.

Blessing, S. B., Gilbert, S. B., Ourada, S., & Ritter, S. (2009). Authoring Model-Tracing Cognitive Tutors. *International Journal of Artificial Intelligence in Education,* 19(2), 189–210.

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.

Brown, J. S., & VanLehn, K. (1980). Repair theory: A generative theory of bugs in procedural skills. *Cognitive Science,* 4(4), 379–426.

Burton, R. R. (1982). Diagnosing bugs in a simple procedural skill. In D. Sleeman & J. S. Brown (Eds.), *Intelligent Tutoring Systems* (pp. 157–183). London: Academic Press.

Burton, R. R., & Brown, J. S. (1976). *A tutoring and student modelling paradigm for gaming environments*. Paper presented at the SIGCSE-SIGCUE Joint Symposium on Computer Science Education.

Callaway, C. B., Dzikovska, M., Farrow, E., Marques-Pita, M., Matheson, C., & Moore, J. D. (2007). *The Beetle and BeeDiff Tutoring Systems.* Paper presented at the 2007 Workshop on Spoken Language Technology for Education.

Carbonell, J. (1970). AI in CAI: An artificial-intelligence approach to computer-assisted instruction. *IEEE Transactions on Man-Machine Systems,* 11(4), 190–202.

Carr, B. P. B. (1977). *Wusor II: A Computer Aided Instruction Program with Student Modelling Capabilities*. Unpublished Master's Thesis, Massachusetts Institute of Technology, Cambridge, MA.

Castellano, M., Mastronardi, G., Di Giuseppe, G., & Dicensi, V. (2007). *Neural Techniques to Improve the Formative Evaluation Procedure in Intelligent Tutoring Systems.* Paper presented at the IEEE International Conference on Computational Intelligence for Measurement Systems and Applications.

REFERENCES (CONTINUED)

Cha, H. J., Kim, Y. S., Park, S. H., Yoon, T. B., Jung, Y. M., & Lee, J. (2006). *Learning styles diagnosis based on user interface behaviors for the customization of learning interfaces in an intelligent tutoring system.* Paper presented at the 8th International Conference on Intelligent Tutoring Systems.

Charniak, E. (1991). Bayesian networks without tears. *AI magazine,* 12(4), 50–63.

Conati, C., Gertner, A., & VanLehn, K. (2002). Using Bayesian Networks to Manage Uncertainty in Student Modeling. *User Modeling and User-Adapted Interaction,* 12(4), 371–417.

Conati, C., & Zhou, X. (2004). *Building and evaluating an intelligent pedagogical agent to improve the effectiveness of an educational game.* Paper presented at the 9th International Conference on Intelligent User Interfaces.

Corbett, A. T. (2001). *Cognitive Computer Tutors: Solving the Two-Sigma Problem*. Paper presented at the 8th International Conference on User Modeling.

Corbett, A. T., & Anderson, J. R. (1995). *Knowledge decomposition and subgoal reification in the ACT programming tutor.* Paper presented at the 7th World Conference on Artificial Intelligence in Education.

Corbett, A. T., & Anderson, J. R. (2001). *Locus of feedback control in computer-based tutoring: impact on learning rate, achievement and attitudes.* Paper presented at the SIGCHI conference on Human Factors in Computing Systems.

Folsom-Kovarik, J. T., Schatz, S., & Nicholson, D. (2010). *Plan Ahead: Pricing ITS Learner Models.* Paper presented at the 19th Conference on Behavior Representation in Modeling and Simulation.

Graesser, A. C., Jackson, G. T., Mathews, E. C., Mitchell, H. H., Olney, A., Ventura, M., et al. (2003). *Why/AutoTutor: A test of learning gains from a physics tutor with natural language dialog.* Paper presented at the 25th Annual Meeting of the Cognitive Science Society.

Graesser, A. C., Lu, S., Jackson, G. T., Mitchell, H. H., Ventura, M., Olney, A., et al. (2004). AutoTutor: A tutor with dialogue in natural language. *Behavior Research Methods,* 36(2), 180–192.

Graesser, A. C., Wiemer-Hastings, K., Wiemer-Hastings, P., & Kreuz, R. (1999). AutoTutor: A simulation of a human tutor. *Journal of Cognitive Systems Research,* 1, 35–51.

Hatzilygeroudis, I., & Prentzas, J. (2004). Using a hybrid rule-based approach in developing an intelligent tutoring system with knowledge acquisition and update capabilities. Expert *Systems with Applications*, 26(4), 477–492.

Heffernan, N. T., Turner, T. E., Lourenco, A. L. N., Macasek, M. A., Nuzzo-Jones, G., & Koedinger, K. R. (2006). *The ASSISTment builder: Towards an analysis of cost effectiveness of ITS creation.* Paper presented at the 19th International FLAIRS Conference.

Hockenberry, M. (2005). *Simple tutors for hard problems: understanding the role of pseudo-tutors.* Paper presented at the SIGCHI conference on Human Factors in Computing Systems.

Johnson, B. G., Phillips, F., & Chase, L. G. (2009). An intelligent tutoring system for the accounting cycle: Enhancing textbook homework with artificial intelligence. *Journal of Accounting Education,* 30–39.

Johnson, W. L. (1990). Understanding and debugging novice programs. *Artificial Intelligence,* 42, 51–97.

Kass, A., Burke, R., Blevis, E., & Williamson, M. (1994). Constructing learning environments for complex social skills. *Journal of the Learning Sciences,* 3(4), 387–427.

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating Training Programs: The Four Levels.* San Francisco, CA: Berrett-Koehler.

Kodaganallur, V., Weitz, R. R., & Rosenthal, D. (2006). An Assessment of Constraint-Based Tutors: A Response to Mitrovic and Ohlsson's Critique of 'A Comparison of Model-Tracing and Constraint-Based Intelligent Tutoring Paradigms'. *International Journal of Artificial Intelligence in Education,* 16(3), 291–321.

REFERENCES (CONTINUED)

Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with Cognitive Tutors. *Educational Psychology Review,* 19(3), 239–264.

Koedinger, K. R., Aleven, V., Heffernan, N., McLaren, B., & Hockenberry, M. (2004). *Opening the door to non-programmers: authoring intelligent tutor behavior by demonstration.* Paper presented at the 7th International Conference on Intelligent Tutoring Systems.

Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education,* 8(1), 30–47.

Lee, C. S. (2007). Diagnostic, predictive and compositional modeling with data mining in integrated learning environments. *Computers & Education,* 49(3), 562–580.

Lesgold, A., Lajoie, S., Bunzo, M., & Eggan, G. (1988). SHERLOCK: A coached practice environment for an electronics troubleshooting job. In J. Larkin, R. Chabay & C. Scheftic (Eds.), *Computer assisted instruction and intelligent tutoring systems: Establishing communication and collaboration* (pp. 49–80). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lynch, C., Ashley, K., Aleven, V., & Pinkwart, N. (2006). *Defining "ill-defined domains"; A literature survey.* Paper presented at the Workshop on Intelligent Tutoring Systems for Ill-Defined Domains at the 8th International Conference on Intelligent Tutoring Systems.

Martin, B., & Mitrovic, A. (2002). *Automatic problem generation in constraint-based tutors.* Paper presented at the 6th International Conference on Intelligent Tutoring Systems.

Martin, B., Mitrovic, A., & Suraweera, P. (2008). ITS domain modelling with ontology. *Journal of Universal Computer Science*, 14(17), 2758–2776.

McQuiggan, S. W., Mott, B. W., & Lester, J. C. (2008). Modeling self-efficacy in intelligent tutoring systems: An inductive approach. *User Modeling and User-Adapted Interaction,* 18(1), 81–123.

Mendicino, M., Razzaq, L., & Heffernan, N. T. (2009). Comparison of Traditional Homework with Computer Supported Homework. *Journal of Research on Technology in Education,* 41(3), 331–358.

Mitrovic, A., Koedinger, K. R., & Martin, B. (2003). *A comparative analysis of cognitive tutoring and constraint-based modeling*. Paper presented at the 9th International Conference on User Modeling.

Mitrovic, A., & Ohlsson, S. (1999). Evaluation of a constraint-based tutor for a database language. *International Journal of Artificial Intelligence in Education*, 10(3–4), 238–256.

Mitrovic, A., & Weerasinghe, A. (2009). *Revisiting Ill-Definedness and Consequences for ITSs.* Paper presented at the 14th International Conference on Artificial Intelligence in Education.

Murray, T. (2003). An Overview of Intelligent Tutoring System Authoring Tools: Updated Analysis of the State of the Art. In T. Murray, S. Blessing & S. Ainsworth (Eds.), *Authoring tools for advanced technology learning environments: Toward cost-effective adaptive, interactive, and intelligent educational software* (pp. 493–546). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Neches, R., Langley, P., & Klahr, D. (1987). Learning, development, and production systems. In D. Klahr, P. Langley & R. Neches (Eds.), *Production system models of learning and development* (pp. 1–53). Cambridge, MA: MIT Press.

Ohlsson, S. (1994). Constraint-Based Student Modeling. In J. E. Greer & G. McCalla (Eds.), *Student Modelling: The Key to Individualized Knowledge-based Instruction* (pp. 167–189). Berlin / Heidelberg: Springer-Verlag.

Ong, J., & Ramachandran, S. (2005). *An Intelligent Tutoring System Approach to Adaptive Instructional Systems* (Contractor Report No. 2005-09). San Mateo, CA: Stottler Henke Associates, Inc.

Ott, S., Imoto, S., & Miyano, S. (2004). *Finding optimal models for small gene networks*. Paper presented at the Pacific Symposium on Biocomputing.

Payne, S. J., & Squibb, H. R. (1990). Algebra mal-rules and cognitive accounts of error. *Cognitive Science*, 14(3), 445–481.

REFERENCES (CONTINUED)

Razzaq, L., Patvarczki, J., Almeida, S. F., Vartak, M., Feng, M., Heffernan, N. T., et al. (2008). The ASSISTment Builder: Supporting the Life Cycle of ITS Content Creation. *IEEE Transactions on Learning Technologies,* 2(2), 157–166.

Reyes, R. L., & Sison, R. C. (2002). *Using Case-Based Reasoning Approach in Planning Instructional Activities*. Paper presented at the 7th Pacific Rim International Conference on Artificial Intelligence.

Rosé, C. P., Jordan, P., Ringenberg, M., Siler, S., VanLehn, K., & Weinstein, A. (2001). *Interactive conceptual tutoring in Atlas-Andes.* Paper presented at the 10th International Conference on Artificial Intelligence in Education.

Schatz, S., Oakes, C., Folsom-Kovarik, J. T., & Dolletski-Lazar, R. (in press). ITS + SBT: A Review of Operational Situated Tutors. Manuscript submitted for publication.

Schulze, R. (2004). *Meta-analysis: A comparison of approaches.* Göttingen, Germany: Hogrefe & Huber.

Shute, V. J. (1990). *Rose garden promises of intelligent tutoring systems: Blossom or thorn?* Paper presented at the 4th Annual Workshop on Space Operations Applications and Research.

Shute, V. J. (1992). Aptitude-treatment interactions and cognitive skill diagnosis. In J. W. Regian & V. J. Shute (Eds.), *Cognitive approaches to automated instruction* (pp. 15–47). Hillsdale, NJ: L. Erlbaum Associates, Inc.

Shute, V. J., & Glaser, R. (1990). A large-scale evaluation of an intelligent discovery world: Smithtown. *Interactive Learning Environments,* 1(1), 51–77.

Shute, V. J., Hansen, E. G., & Almond, R. G. (2008). You Can't Fatten A Hog by Weighing It – Or Can You? Evaluating an Assessment for Learning System Called ACED. *International Journal of Artificial Intelligence in Education,* 18(4), 289–316.

Snow, R. E., & Swanson, J. (1992). Instructional psychology: Aptitude, adaptation, and assessment. *Annual Review of Psychology,* 43(1), 583–626.

Stottler, R. H., Fu, D., Ramachandran, S., & Vinkavich, M. (2001). *Applying a Generic Intelligent Tutoring System (ITS) Authoring Tool to Specific Military Domains.* Paper presented at the Interservice/Industry Training, Simulation & Education Conference.

Suraweera, P., & Mitrovic, A. (2004). An intelligent tutoring system for entity relationship modelling. *International Journal of Artificial Intelligence in Education*, 14(3–4), 375–417.

Suraweera, P., Mitrovic, A., & Martin, B. (2007). *Constraint Authoring System: an empirical evaluation*. Paper presented at the 13th International Conference on Artificial Intelligence in Education.

VanLehn, K. (1982). Bugs are not enough: Empirical studies of bugs, impasses and repairs in procedural skills. *Journal of Mathematical Behavior,* 3(2), 3–71.

VanLehn, K. (1988). Student modeling. In M. C. Polson, J. J. Richardson & E. Soloway (Eds.), *Foundations of intelligent tutoring systems* (pp. 55–78). Hillsdale, NJ: L. Erlbaum Associates.

VanLehn, K., Graesser, A. C., Jackson, G. T., Jordan, P., Olney, A., & Rosé, C. (2007). When Are Tutorial Dialogues More Effective Than Reading? *Cognitive Science*, 31, 3–62.

VanLehn, K., Jordan, P., Rosé, C., Bhembe, D., Böttner, M., Gaydos, A., et al. (2002). The Architecture of Why2-Atlas: A Coach for Qualitative Physics Essay Writing. In S. A. Cerri, G. Gouarderes & F. Paraguacu (Eds.), *Intelligent Tutoring Systems* (pp. 158–167). Berlin / Heidelberg: Springer.

VanLehn, K., Lynch, C., Schulze, K., Shapiro, J. A., Shelby, R., Taylor, L., et al. (2005). The Andes physics tutoring system: Lessons learned. *International Journal of Artificial Intelligence and Education*, 15(3), 147–204.

Walonoski, J. A., & Heffernan, N. T. (2006). *Detection and analysis of off-task gaming behavior in intelligent tutoring systems.* Paper presented at the 8th International Conference on Intelligent Tutoring Systems.

Wang, N., Johnson, W. L., Rizzo, P., Shaw, E., & Mayer, R. E. (2005). *Experimental evaluation of polite interaction tactics for pedagogical agents.* Paper presented at the 10th International Conference on Intelligent User Interfaces.

Understanding the Value of M&S

AUTHORS

Jerry M. Feinberg
Robert J. Graebener
Kriya S. Kaping
Michael J. Meehan

Alion Science and Technology / Modeling and Simulation Information Analysis Center (MSIAC)
1901 N. Beauregard St., Suite 400
Alexandria, VA 22311
709-933-3360

jfeinberg@alionscience.com
rgraebener@alionscience.com
mmeehan@alionscience.com

ABSTRACT

Modeling and simulation (M&S) underpin the development and application of vital capabilities that can be realized in no other way. These capabilities that M&S delivers provide real value to the nation as a whole; to the Department of Defense and other Government departments; to industry; and to scientists and engineers. The bottom line is that people need to get things done, and M&S is the go-to tool for enabling actions.

Efforts at measuring the value of M&S have been sporadic and inconclusive due to a variety of issues. Many approaches have concentrated on measuring M&S activities. However, we believe that decision makers do not care about M&S *per se*, but that they do care about and desperately need the capabilities that M&S can deliver. This paper focuses on the value provided to decision makers by these M&S capabilities.

We have developed and started applying a framework for determining the value of M&S to decision makers. Our goal is to use this framework to communicate an understanding of value to these same decision makers and M&S stakeholders. The framework, based on a hierarchy of decision makers, encompasses M&S usage in four tiers: *society/government*, enterprise, community, and program. We are currently investigating ways to specify the value that M&S provides stakeholders at each of these tiers. Values include the ability to perform, improvements in performance, savings of lives, cost avoidance, return on investment, assurance, reliability, interoperability, re-use, and availability of resources.

This paper begins with an overview of applications of M&S supporting decision makers. It then defines M&S for our purposes, describes the uses of M&S, and defines the value of M&S. This is followed by a specification of a hierarchy of decision makers using M&S and then a discussion of value and the metric *return on investment* (ROI). Next, the paper introduces the framework for the value of M&S, provides examples of decision makers at different tiers within the hierarchy, and introduces the concept of the *M&S Value Network*. The paper concludes by offering many additional examples of M&S usage and value that illustrate the utility of this framework and the importance of articulating M&S value within this framework.

KEYWORDS

Economics of M&S, cost effectiveness, return on investment

ABOUT THE AUTHORS

DR. JERRY FEINBERG is the Chief Scientist for the MSIAC in Alexandria, VA. Associated with the MSIAC from its inception, he leads or guides many of its technical projects. He holds a Ph.D. and M.S. in Mathematics and an M.S. in Physics from Stanford University, and a B.S. in Mathematics from the California Institute of Technology. Dr. Feinberg is a member of Tau Beta Pi and is a Hertz Foundation Fellow.

MR. BOB GRAEBENER is the Director of the MSIAC where he leads a team of professionals responsible for enabling quick access to worldwide scientific and technical information to improve the productivity of personnel in the Defense science and technology communities. He has over thirty years of experience in program and project management, strategic planning and decision-making, M&S, operations research and training. Mr. Graebener holds

an M.S. in Industrial Management from the Georgia Institute of Technology and a B.S. in Industrial Engineering and Operations Research from the Virginia Polytechnic Institute and State University. He is also a Project Management Professional (PMP®).

MR. MIKE MEEHAN is the Help Desk Supervisor for the MSIAC in Alexandria, VA, and is a graduate of the University of South Carolina. He has worked as part of the MSIAC staff since its creation, working initially as part of the integration cadre building the first DoD M&S Repository and most recently on the deployment of the DoD M&S Catalog and the DoD Modeling and Simulation Information System (MSIS). Mr. Meehan has over thirty-five years of experience supporting the DoD in the areas of training, command and control, and program management.

INTRODUCTION

Modeling and Simulation (M&S) is critical for our nation. M&S supports meeting security challenges by enhancing our economic competitiveness, our energy independence, the health of our environment and the health of our citizens, and, of course, our national defense. Some examples of specific capabilities underpinned by M&S include crisis prediction, response to natural disasters, transportation planning, and support to the warfighters and first responders, both in the field and in transit. The contributions of M&S to the medical health of our nation include not only pandemic response, but also support to developing new drugs for treating devastating diseases and developing new prosthetics for returning our wounded warfighters to independence. M&S also extends the reach of our medical experts into dangerous areas of operation and into underserved areas of our own country. In short, M&S saves and repairs lives, saves taxpayer dollars, and improves our operational capabilities.

M&S delivers real value to the nation as a whole, to the Department of Defense and other Government departments, to industry, and to scientists and engineers. The bottom line is that people need to get things done, and M&S is the go-to tool for enabling actions. While the use of M&S is ubiquitous, there have been few systematic investigations into ways to describe the value M&S provides. Efforts at describing this value have been sporadic and inconclusive due to a variety of issues. Although decision makers may neither care about nor acknowledge M&S tools *per se,* they do care about (and desperately need) the capabilities that M&S tools can deliver together with an appreciation of *how* these capabilities contribute to solving our most critical national challenges. That is, decision makers need to understand M&S value.

Consequently, the MSIAC has initiated an effort to understand the value of M&S and to communicate this value to decision makers and M&S stakeholders. As a first step, we have developed a framework based on tiers of decision makers who rely on M&S. This framework stratifies the values that decision makers receive from M&S and permits investigators to "drill down" into specific sets of users. Section 2 starts with a definition of M&S for our purposes, describes the uses of M&S, and then defines the value of M&S. Next, Section 3 presents the hierarchy of decision makers using M&S. Section 4 discusses M&S value, details one important metric – return on investment (ROI), and then indicates why ROI should not be considered as the universal metric for M&S. Following this, Sections 5 and 6 introduce the framework for the value of M&S and provide examples of value at different tiers within the framework's hierarchy. Section 7 introduces the concept of the *M&S Value Network*. Finally, Section 8 provides additional examples of M&S usage and value that indicate the utility of this framework approach.

M&S, USE OF M&S, AND THE VALUE OF M&S

We define M&S as "an organized approach for investigating, interpreting, understanding, and practicing real-world behaviors, situations, and processes." M&S can represent current or anticipated systems, people, processes, and environments in a consistent fashion for developing insights, specifications, predictions, and skills to improve design, construction, training, or operation. M&S is used to support decision makers by:

- elucidating results,
- envisioning effectiveness,
- enabling actions,
- enhancing capabilities, or
- evoking/replicating real experiences.

As in economics, politics, and sociology, we consider "value" to be a measurement that assesses the degree to which something meets the needs or influences the decision making of a group. Consequently, for defining the value of M&S, we follow [1] and define value as "the worth of M&S based on the wants and needs of the members of society." This approach corresponds to the subjective

theory of value in economics where value is not considered to be something intrinsic to a particular application of a tool, but is accrued only because that tool provides something of worth to the society that uses it. Consequently, we are not concerned with "activity metrics" (the number of times something occurs) but instead we will attempt to account for the value of M&S through the worth of the results that it supports at the tier(s) of society under examination.

These tiers of society are the basis for a hierarchy of M&S users that is presented next.

HIERARCHY OF DECISION MAKERS USING M&S

Since M&S is both pervasive and ubiquitous, we believe that it is unreasonable to consider M&S users to be a homogeneous bloc. Decision makers operate in differing environments, they are very specialized and focus on their specific issues, and these issues vary from one decision-maker to another. Because of this, we introduce a hierarchy of users who rely on M&S to support decisions and use this hierarchy to further our investigation into understanding the value of M&S. This stratification is adopted from methodology developed in [2] and is analogous to the usual partitioning of models into four tiers corresponding to their scope: campaign models, mission models, engagement models, and system models (Figure 1).

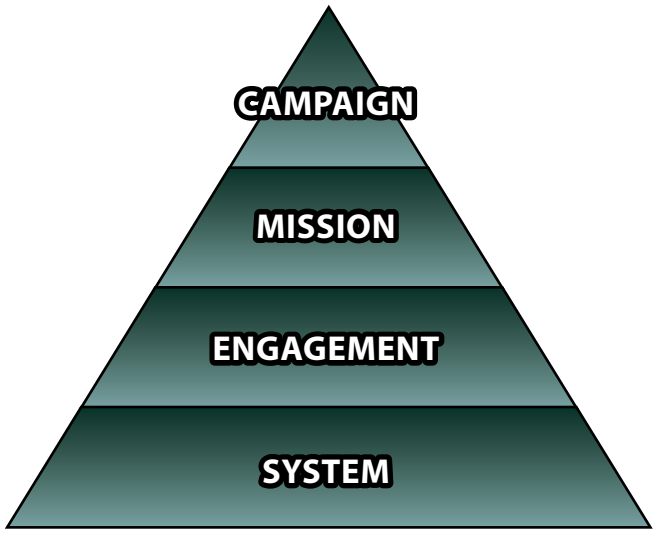


Figure 1 – Hierarchy of Models

Specifically, we stratify decision makers supported by M&S into four tiers (see Figure 2). At the highest level, M&S supports decisions at the *society/government* tier. The decision makers here include members of the U.S. Government Executive Branch and members of Congress. The enterprise tier is the second level. For the DoD, the decision makers at this tier include the leaders of the Combatant Commands (COCOMs), Joint warfighters, and the Service M&S office directors (e.g., Army Modeling and Simulation Office – AMSO, Navy Modeling and Simulation Office – NMSO, Air Force Agency for Modeling and Simulation – AFAMS). Below this is the community tier. This tier is introduced specifically to account for the decision makers in the seven "communities enabled by M&S" that comprise a portion of the DoD M&S governance structure [3].

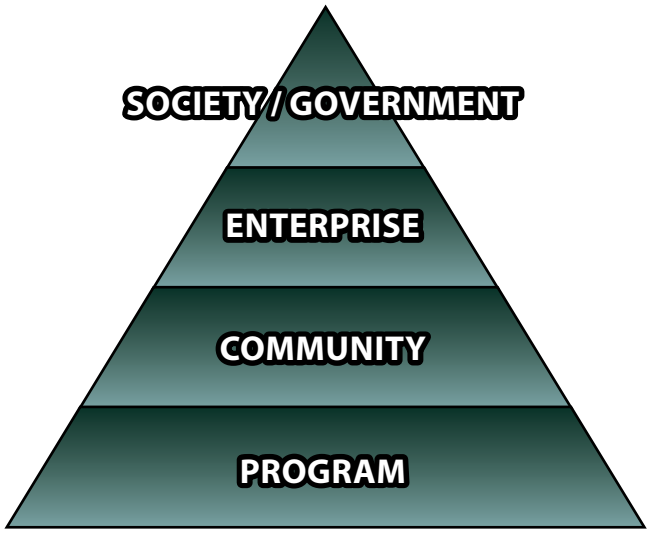


Figure 2 – Hierarchy of Decision Makers

At the fourth level, M&S supports decisions at the program tier. This contains much of where the "rubber" of M&S meets the "road" of program applications and includes the decision makers in individual acquisition programs such as Joint Strike Fighter (JSF), specific M&S programs such as Joint Integrated Contingency Model (JICM) and Synthetic Theater Operations Research Model (STORM), most of the industry and academic research and development (R&D) efforts supporting enhancements to M&S, and the manufacturing community (both hardware and software). Examples of decision makers at these varying tiers are provided in Table 1.

DECISION-MAKER TIER	EXAMPLES OF DECISION MAKERS AT TIER
Society / Government	• Legislative Branch (Senate, House of Representatives) • Federal Agencies (DoD, NASA, DOE, DHS, …) • Other Government (State, Local, Regional, …) • Organizations (SISO, SCSI, IEEE, …), Academia, Industry (Medicine, Manufacturing) • Foreign Governments • Allied Coalitions
Enterprise	• COCOMs • Senior OSD Decision Makers • Service Office Directors, Leaders, Champions for Service M&S
Community	• Leaders of M&S-Enabled Communities (Currently: Acquisition, Analysis, Experimentation, Intelligence, Planning, Testing, Training)
Program	• Simulation Developers • Acquisition Program Managers (JSF, DD-1000, …) • Simulation Program Managers (JICM, STORM, ITEM, NSS, OneSAF, …) • M&S Tools, Data, Services (DoD M&S Catalog, Environmental Data Cube Support System, Rapid Data Generation, …)

Table 1 – Examples of Decision Makers at Varying Decision-Maker Tiers

We note that there are not clean separations between the various tiers of decision makers in the hierarchy. That is, individual decision makers could easily be placed in two (or even three) of the tiers. This matches the situation for the modeling hierarchy (Figure 1), where the models there can also span multiple tiers. We also note that one could easily propose alternate hierarchies differing in the number of tiers or in the definitions for the tiers. However, our major point in this paper is the absolute need to stratify the decision makers so that what is important about M&S to each of them can be identified and then measured in the decision maker's own context. We believe that the primary utility of this hierarchy (and the framework described below) is to organize our thoughts for preparing to tackle the real issues of determining values for the major stakeholder groups, without either omitting inputs or missing opportunities to expand on identified worth.

The next section discusses these values.

VALUE AND THE RETURN ON INVESTMENT (ROI)

We have equated the value of M&S to the worth of the results that it supports. Our next step is to describe the value supplied by M&S to decision makers. Some possible values for the capabilities provided by M&S include, but are certainly not limited to:

- the ability to perform,
- improvements in performance,
- savings of lives,
- cost avoidance,
- return on investment,
- assurance,
- reliability,
- interoperability,
- re-use, and
- availability of resources.

Note that none of these values applies across the entire hierarchy of decision makers. As will be presented in Section 6, some of these values are more suited for application at the higher tiers, and others at the lower tiers.

One commonly applied metric for measuring the value of M&S (and of other technologies/programs) is *Return on Investment,* or ROI. There are many essentially equivalent definitions for ROI, but an interesting one is given in the DoD glossary for training [4], which states that ROI is "the savings that accrue, usually over a predetermined period of time, either through cost avoidance or savings in time, money, personnel, or other resources, by an up-front investment in a … program." We note the explicit recognition of cost avoidance in this definition since many other definitions of ROI focus exclusively on monetizing all the returns and consider only money savings. However, even our broader definition does not include all of the aspects of value or worth that we argue must be considered, especially at the higher decision-maker tiers.

There are additional issues to consider when investigating the applicability of ROI for government organizations [2]. Government accounting characteristics are different than commercial business practices; they vary in areas such as the lack of profit incentive, distribution of bill-payers

and value recipients, and intangible value recovery over indefinite timeframes. Consequently, we believe that even though ROI can be a useful metric for value in specific situations, we believe that ROI cannot be used as the single metric for measuring value at all tiers.

The following illustrates and clarifies some of the differences between value at the higher tiers and ROI. The basic scenario is a contemplated purchase of a routing/planning system for your car. The available systems are either a standalone GPS system or an equivalent app for your smartphone (both based on M&S!). So, (1) how do you decide whether or not to buy a system (at all), and (2) how do you decide which approach, brand or model is best for you? We believe that you make the basic decision of whether or not to buy based on the worth (value) to you in being able to plan routes and estimate arrival times efficiently. Since these M&S-based systems are the only way that you will be able to perform these functions (outside of having permanent reach-back to an automobile association or of hiring a full time "shotgun" who is not only adept at finding, unfolding, reading, and re-folding maps, but also can perform amazing real-time mental calculations), your decision to purchase is based on your need to perform, and the performance is M&S value. Granted, cost should be considered, at least implicitly, since you will (eventually) have to pay the bill for the system. Now that the fundamental value – ability to perform – has been specified, you can make the choice of which brand of GPS system or smartphone app to buy. This choice can be supported by considering ROI where the *return* (value) involves an estimate that combines the worth of the units' specific features and their ease of use, and the *investment* is the actual amount you will pay to obtain these features from the differing systems.

The bottom line is that the value of the M&S-based system accrues from the capability to perform, while the concept of ROI enters later in making a specific choice.

The next section introduces the framework combining the hierarchy of decision makers and the corresponding values.

THE FRAMEWORK FOR THE VALUE OF M&S

The analysis of the tiers of decision makers presented in Section 3 indicates that it is unrealistic to consider the decision makers who rely on M&S to be a homogeneous bloc. Similarly, Section 4 shows that it is unrealistic to consider the worth these decision makers derive from M&S to be represented by a single value. This perspective suggests a framework designed specifically to account for the differentiation of values. This simple framework, presented in Table 2, explicitly recognizes the tiers corresponding to the decision-maker hierarchy and separates the values provided by M&S to the decision makers at each tier. We believe that this approach is the key both to advancing the understanding of M&S value and to communicating this understanding to the decision makers themselves.

DECISION MAKER TIER	VALUES OF M&S AT DECISION-MAKER TIER
Society / Government	Value of M&S at the *Society / Government* Tier
Enterprise	Value of M&S at the *Enterprise* Tier
Community	Value of M&S at the *Community* Tier
Program	Value of M&S at the *Program* Tier

Table 2 – The Value of M&S Framework

The next section supplies some explicit examples of values at the various tiers.

EXAMPLES OF VALUE AT DIFFERENT TIERS IN THE HIERARCHY

This section adds content to the framework introduced above for measuring the value of M&S. Table 3 presents suggested examples of values for different tiers of decision makers. As shown in the table, there are significant differences in the values of M&S at the different decision maker tiers. (Of course, that is the reason for building the framework and hierarchy in the first place.) Since the society/ government decision makers tend to concentrate on making go/no-go decisions and reducing risk and uncertainty, the values of M&S they use include the basic ability to perform a mission or operation, the ability to perform the mission or operation within critical time constraints, the ability to improve the performance of a mission or operation, the reduction of risk, the reduction of uncertainty, and the savings of lives.

DECISION MAKER TIER	EXAMPLES OF VALUE OF M&S AT DECISION-MAKER TIER
Society / Government	• Basic Ability to Perform a Mission • Ability to Perform a Mission Within Time Constraints • Ability to Perform Better • Reduction of Risk • Reduction of Uncertainty • Savings of Lives • Reduction of Cost
Enterprise	• Ability to Perform in Limited Time • Ability to Perform Better • Reduction of Risk • Reduction of Uncertainty • Improved Interoperability • Reduction of Cost • Savings of Lives • Improved Re-use • Reduced Duplication
Community	• Reduction of Risk • Reduction of Uncertainty • Ability to Perform Better • Improved Interoperability • Reduction of Cost • Cost Effectiveness • Return on Investment • Improved Re-use • Reduced Duplication
Program	• Reduction of Risk • Reduction of Uncertainty • Ability to Perform Better • Savings of Lives • Improved Interoperability • Reduction of Cost • Return on Investment • Cost Effectiveness • Improved "ilities"

Table 3 – Examples of Value at Varying Decision- Maker Tiers

At the enterprise tier, the values for M&S also include improvements in interoperability and reductions in cost. At the community and program tiers, values for M&S relating to cost, cost effectiveness, and ROI are important. Finally, at the program tier, the values for M&S also include improving the usual "ilities" such as maintainability and reliability.

THE M&S VALUE NETWORK

In the modeling hierarchy (Figure 1), the inputs to models in a higher tier are usually based on the measures of effectiveness output by elements in the lower tiers. We believe that an analogous data flow occurs between the tiers of the framework for value of M&S: values at lower tiers support the values at higher tiers. For example, the values of uncertainty reduction and risk reduction are supported by the values of ilities such as reliability and maintainability. Similarly, at the higher tiers, the value of the basic ability to perform is supported by the value of improved interoperability.

This flow of values, together with the values within the tiers, forms what we refer to as the *M&S Value Network* (see Figure 3). As we continue expanding our effort reported in this paper, we expect to develop more details that illustrate this M&S value network. We also anticipate that the study of the value interactions forming this network may be as important as the study of the values themselves, as it is the network that indicates how the members of each tier communicate their perceived value of M&S to those of the other tiers.

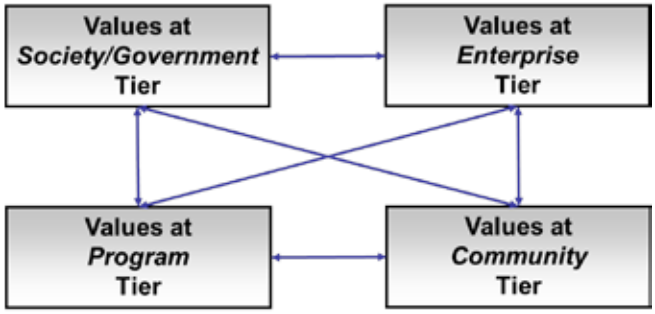


Figure 3 – M&S Value Network

ADDITIONAL EXAMPLES OF M&S VALUE

This section offers additional examples illustrating uses of M&S by decision makers in the tiers of the framework hierarchy. By taking these examples primarily from documents or statements produced by decision makers themselves, this section offers additional insight into the capabilities of M&S to provide value. These examples focus at the top tier and indicate that a large part of M&S value corresponds to enabling capabilities that can be realized realistically only through applying M&S.

These examples also indicate the need to communicate these values: the ability to articulate the value of M&S to society is just as necessary as determining its value. For example, in the *society/government* tier of decision makers, Congressional members need to communicate the value of M&S to their constituents and fellow members of government to maintain required funding for M&S research and development. Similarly, members of Congress have challenged the DoD M&S community to communicate in simple terms the importance of M&S as a nationally critical technology so they can continue their support. This communication must also demonstrate to the average person *how* M&S technologies solve our Nation's major issues, rather than dwelling on the technical details of *what* comprises an M&S technology.

For DoD (*society/government* tier), [5] states that "DoD has made a firm commitment to leverage M&S by planning and implementing a broad range of initiatives that apply this national critical technology to saving lives, preserving taxpayer dollars, and increasing operational capabilities." The values corresponding to these M&S capabilities include the basic ability to perform a mission, ability to perform a mission within time constraints, ability to perform better, reduction of risk, reduction of uncertainty, and savings of lives.

The same document [5] continues with "the Department cannot function effectively without M&S; from the commanders in the field, through the acquisition executives, to the research and development institutions, the Department depends on capabilities, insights, and understandings enabled by M&S." Again, the values for these *society/government* tier decision makers encompass all of those noted immediately above.

Also from [5], "M&S enables strategic operations and support functions to our military, aids plans for national disaster responses and emergency preparedness, fosters and maintains our strategic partnerships, and enhances global economic competitiveness." The values for these top-tier decision makers include the basic ability to perform a mission, ability to perform better, reduction of risk, and savings of lives.

For a member of the U.S. Congress (*society/ government* tier), a value of M&S is its "capability to support good, high-paying jobs in a member's district" [6]. Clearly, there are alternatives for obtaining this value, so ROI might be a good fit for this situation. We note that at this tier there is also a great need to communicate the value of M&S to further investments and advancements.

From [7], "the massive economic recovery package was designed largely on the basis of predictions developed using products of the M&S industry." The values here for decisions made at the *society/government* tier include the basic ability to perform a mission, ability to perform better, reduction of risk, and reduction of uncertainty.

Again from [7], "the Department of Energy (DOE) has initiated the Advanced Simulation & Computing (ASC) campaign, using high-performance simulations to tackle a wide variety of scientific issues such as the safety of the nuclear stockpile and the development and improvement of alternative energy sources. This vital issue involves decision makers at the *society/government*, *enterprise*, and *program* tiers. The values directly include the basic ability to perform a mission, reduction of risk, and reduction of uncertainty. Other values could include cost, cost effectiveness, and return on investment.

Again from [5], M&S provides "Combatant Commanders and supporting agencies with the ability to understand complex interactions, to apply emerging technological capabilities as force multipliers, and to imagine the yet-to-be-imagined for providing innovative solutions to meet national security challenges." The values for these *enterprise* tier decision makers include the ability to perform in limited time, ability to perform better, reduction of risk, and reduction of uncertainty.

For the Service components of a Joint task force (*enterprise* tier), decision makers use M&S capabilities to "synchronize, learn, and practice or train more effective tactics, techniques and procedures for applying complex systems as force multipliers to succeed on the battlefield against a man-made event or during support relief to a natural disaster or incident" [5]. The values here include the ability to perform better, reduction of risk, reduction of uncertainty, and improved interoperability.

Once again from [5], "the Department uses M&S to provide insights and predictions for the performance of current or anticipated systems for improving design and construction, training, and operations." These M&S capabilities are at the *program* tier and the corresponding values include reduction of risk, reduction of uncertainty, ability to

perform better, improved interoperability, reduction of cost, return on investment, cost effectiveness, and better "ilities."

For planners at the *program* tier, decision makers use M&S to support "cross-organization collaboration" and the capability to "broaden the scope of responses without the prohibitive outlay of resources that a non-M&S approach would require" [5]. The values include the ability to perform better, improved interoperability, reduction of cost, and cost effectiveness.

From [7], "the most common application of M&S in the Justice system is in the area of accident reconstruction. The development of M&S applications to study incidents in the transportation sector has become quite common. For significant events, such applications of M&S are essentially mandatory to demonstrate the sequence of events and to help establish cause." The decision makers in the Justice system cited here span several of the tiers of the framework and the values include the basic ability to perform a mission, ability to perform better, reduction of risk, and the reduction of uncertainty. We also suggest that additional values should include reduction of cost, cost effectiveness, return on investment, and better "ilities."

We end this section with two additional examples at the *enterprise* tier, namely weather forecasting and the insurance industry. Both of these endeavors are wedded to M&S technology primarily because there are no other viable approaches to support their effectiveness.

In summary, these examples indicate that the values at the various tiers differ from one another, and that the values at the highest *society/government* tier in the framework are more heavily involved with the basic ability to perform a mission or an operation, and not with ROI.

NEXT STEPS

M&S is an enabler to our country in many ways. This paper has presented the first steps in our effort for understanding M&S value. Our goal is to identify and elucidate:

- who is the recipient of M&S value,
- what is the value added by M&S to the recipient,
- why does M&S add value,
- how does M&S add value, and
- where does M&S add value?

We anticipate that our results, based on the framework described in this paper, will mitigate many of the existing problems that have surfaced in previous attempts to quantify the value of M&S. This paper has proposed a workable framework, provided examples of decision makers at the various tiers, and provided examples of M&S values at these tiers. Our next steps are to continue refining and populating this framework with M&S values that are meaningful, can be measured, and will communicate the essentialness of M&S. Future papers will also describe the M&S value network – interactions between tiers that support the measurements of M&S value, and the role of outreach in promoting M&S and enhancing its value.

REFERENCES

[1] Menger, Carl: Principles of Economics. Ludwig von Mises Institute, 2007

[2] Metrics for Modeling and Simulation (M&S) Investments - Scientific and Technical Report November 2008. The AEgis Technologies Group, Inc.

[3] DoD Modeling & Simulation Management System: M&S Coordination Office (M&SCO). CD-ROM, December 2010

[4] MIL-HDBK-29612-4A: Department Of Defense Handbook, Glossary for Training (Part 4 of 5 Parts). 31 August 2001

[5] Report on Department of Defense Joint Modeling and Simulation Activities. Office of the Director, Defense Research & Engineering, January 2009

[6] Remark by a Member of Congress at the Press Conference Announcing the Bill to Establish Grants for the Study of M&S at Institutions of Higher Learning. VMASC: August 19, 2008.

[7] Proposal to Add a North American Industry Classification System (NAICS) Code for the Modeling and Simulation Industry. Prepared and Submitted 22 April 2009 by: National Training and Simulation Association et al. 2010

Permission is hereby granted to quote any of the material herein, or to make copies thereof, for non-commercial purposes, as long as proper attribution is made and this copyright notice is included. All other uses are prohibited without written permission from SISO, Inc.

ABOUT THE M&S JOURNAL

THE M&S JOURNAL is a theme-based, quarterly publication of articles that highlight M&S technology, applications, prototype processes or products, points of view, or emerging philosophies. The M&S Journal is a valuable resource for the M&S community: across DoD, other government agencies, international partner organizations, industry, and academia. All submitted technical papers for the M&S Journal undergo rigorous peer review following an initial screening for conformance to basic requirements. Publishing in the M&S Journal affords authors both an online and print forum for their M&S technical papers, gaining recognition and publicity throughout the DoD M&S Community. Authors also receive extended visibility for their ideas through free online access to their article on the Modeling and Simulation Coordination Office (M&SCO) website. **http://msco.mil**

SUBSCRIPTIONS: If you would like to subscribe to the M&S Journal, please send an email to:

journal-subscribe@lists.dod-msiac.org.

The appearance of an article in the M&S Journal does not constitute an endorsement by the DoD, the Modeling and Simulation Coordination Office (M&SCO), the Defense Technical Information Center (DTIC), the Modeling and Simulation Information Analysis Center (MSIAC), or any of the affiliated government contractors.

Reproduction of non-DoD articles is subject to original copyright restrictions.

Distribution Statement A: Approved for public release: distribution unlimited.

INVITATION TO SUBMIT AN ARTICLE

AUTHOR GUIDELINES

- Submissions may be entirely new work, or previously published papers that would benefit from a wider exposure and would provide valuable resources for M&S users.
- Submission must be previously cleared material for open distribution, and must include a reprint permission statement.
- Manuscripts should be between five to fifteen pages, or 500 to 5,000 words.
- Manuscripts should be submitted in Microsoft Word format.
- The M&S Journal Editorial Board reserves the right to modify a paper for the purpose of typographical or grammar corrections.
- The author will be notified whether the submission has met the basic requirements for the M&S Journal, and will be notified again when the final acceptance/ rejection decision has been made.
- The M&S Journal does not accept papers that are structured as commercial advertising, or as promotions of products or services.

If you would like to submit a technical paper for consideration to the Journal, please e-mail it together with contact information to:

MsiacHelpDesk@DoD-MSIAC.org.

Notes

EDITORIAL BOARD

GUEST EDITOR

ROGER S. SAMUELS — *Chief, Simulation Proponent & School Headquarters, Department of the Army Deputy Chief of Staff G8, Center for Army Analysis*

EDITORIAL BOARD

DR. J. DAVID LASHLEE — *Associate Director for Data, Modeling and Simulation Coordination Office (M&SCO)*

MR. ROBERT GRAEBENER — *Director, Modeling and Simulation Information Analysis Center (MSIAC)*

DR. JERRY FEINBERG — *Chief Scientist, Modeling and Simulation Information Analysis Center (MSIAC)*

MS. SARNA MARCUS — *Marketing and Outreach Manager, Modeling and Simulation Information Analysis Center (MSIAC)*

PUBLICATION STAFF

MR. CHRISTOPHER ELLIS — *Project Manager, Marketing and Outreach Specialist, Modeling and Simulation Information Analysis Center (MSIAC)*

MR. LANGDON GAGNE — *Senior Graphic Designer, Alion Science and Technology*

MS. KELLIE PINEL — *Publications Editor, Modeling and Simulation Information Analysis Center (MSIAC)*

MS. SHANNON REDWINE — *Publications Editor, Modeling and Simulation Information Analysis Center (MSIAC)*

WINTER EDITION 2011

THE M&S JOURNAL

VOLUME 6 ISSUE 3

A PUBLICATION OF
THE MODELING & SIMULATION COORDINATION OFFICE
(M&SCO)