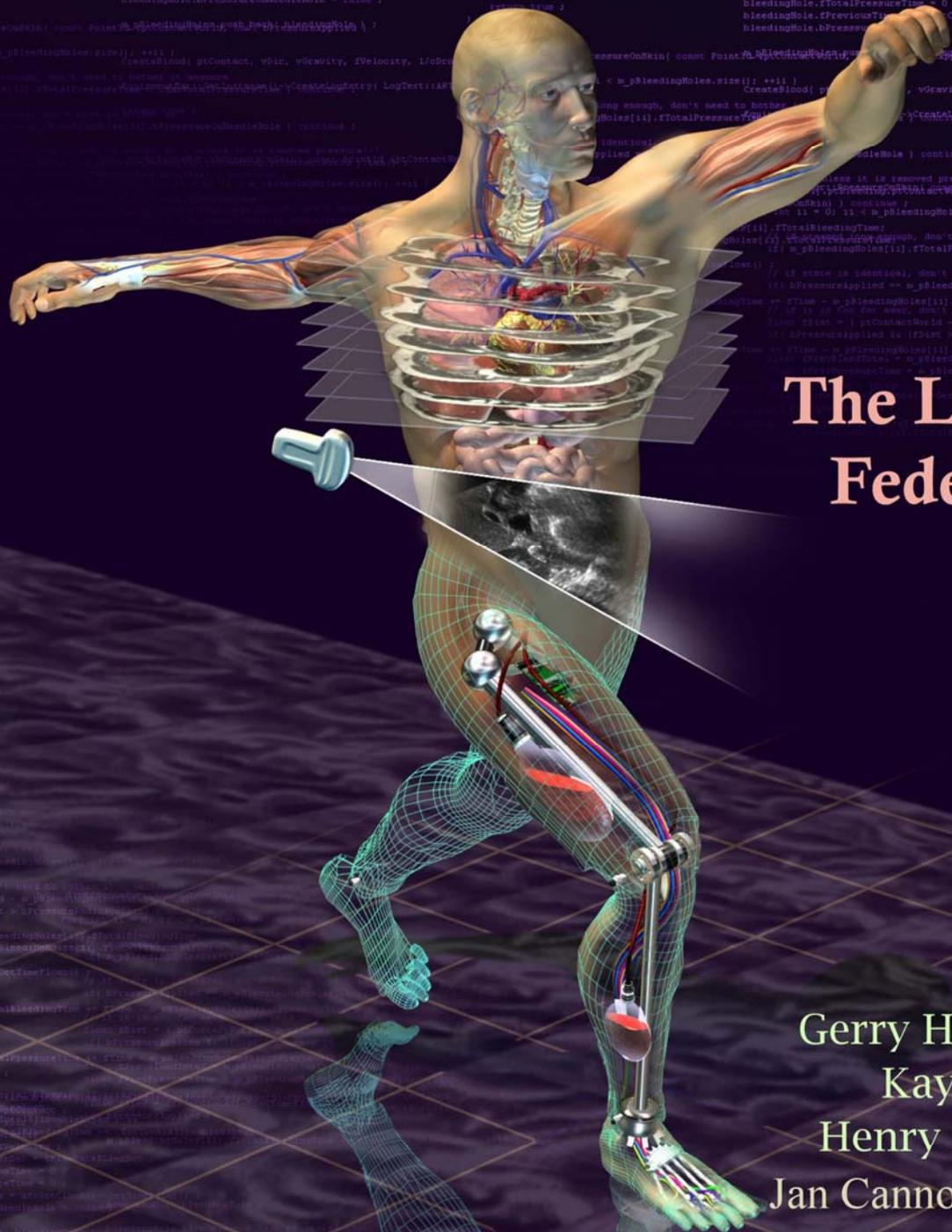


VIRTUAL PATIENT RESEARCH ROADMAP



The Learning Federation

Gerry Higgins, Ph.D
Kay Howell
Henry Kelly, Ph.D
Jan Cannon-Bowers, Ph.D.



Laerdal
helping save lives



The Virtual Patient R&D Roadmap incorporates work from a series of technology research roadmaps, the Learning Science and Technology R&D Roadmap, developed over a three year period by the Federation of American Scientists and the Learning Federation, a partnership among industry, academia, and private foundations to stimulate research and development in learning science and technology. The full series of research roadmaps is available at www.fas.org/thelearningfederation.

The Virtual Patient

A Research Roadmap for the Integration of Learning Technologies into Medical Education

TABLE OF CONTENTS

Overview	1
1 Introduction	2
2 A Brief History of Simulation-Based Training in Medicine	3
3 Relevant Findings from Flight and Defense Simulation Training	9
4 Key Concepts from Learning Science	14
5 Research Challenges and Topics	21
5.1 Research Task #1: Automated generation of patient case/scenarios	24
5.2 Research Task #2: Modeling learners and intelligent tutoring	37
5.3 Research Tasks #3: Feedback and assessment	43
6 Summary and Implications of the Virtual Patient Roadmap	48
6.1 Curriculum Redesign to Integrate New Learning Technologies	48
6.2 Growth in Simulation Centers	49
6.3 Integration of Learning Sciences into the Medical Simulation Curriculum	51
7 References	53
8 Virtual Patient Workshop Participants	64
9 Virtual Patient Workshop Agenda (June 2005)	67

•
•
•
•
•



This page intentionally blank

The Virtual Patient

A Research Roadmap for the Integration of Learning Technologies into Medical Education

Overview

This document presents a research and development plan, or “road map,” designed to improve the integration of learning technologies into simulation-based trainers in medicine. The goal is to form an effective bridge between textbook and patient, while reducing errors associated with the acquisition of patient care skills. This research road map fills a critical need to raise awareness of research challenges and R&D priorities for next-generation medical simulators.

Stakeholders need to have a coordinated understanding of the relevant research results, computational tools, on-going programs and projects across research disciplines, industry efforts, and government funding organizations. This roadmap will hopefully encourage dialog and partnerships to leverage gains from one field to other fields. The road map will provide the non-medical learning technologies community background in the past, current and future of patient simulation. It will also serve as a central resource on medical simulation for the community of practitioners, educators and technology developers.

The road map incorporates the results of a workshop held June 27 -28, 2005 at the University of Maryland School of Medicine attended by fifty participants, representing the fields of allied health and medical education and training, information technology, and the learning sciences. The workshop identified 3 key focus research topics and tasks, milestones, and performance measures for each of the research topics. These were further refined via digital collaborative tools. Many of the workshop participants contributed to the writing of this document, which is truly the product of the community of clinicians, technologists, and learning scientists who participated in the workshop.

1 Introduction

With increasing burdens placed on medical colleges and schools of allied health to enhance medical education to improve healthcare delivery, renewed emphasis has been placed on new approaches for more efficient and effective learning of cognitive and psychomotor skills in medicine. Computer-based simulation educational and training systems have become increasingly useful in medical education. These systems range from manikins to virtual reality environments and can track a student's progress, provide feedback and remediation, and provide practice opportunities without harming actual patients.

However, as took place with flight simulators over the past 60 years, we need to establish and disseminate best practices for use of patient simulators and other computer-based simulation tools throughout medical education and training (from pre-hospital to postgraduate). Although medical education is based on mentorship, the manpower demands for any particular senior medical practitioner, be they master surgeon or chief EMS educator; inhibit individualized contact which is critical for successful learning outcomes.

Research from recent studies in learning technologies show that certain computer-based strategies can enhance learning, employing factors such as feedback, real-time assessment, repetitive practice and intelligent tutoring. This roadmap draws from a workshop sponsored by The Learning Federation called 'The Virtual Patient' that was held in 2005, involving medical simulation experts, game developers, learning scientists, and focusing on three specific domains:

- Automated Generation of Patient Case Scenarios
- Learner Modeling and Intelligent Tutoring
- Feedback and Evaluation

In all areas, the emphasis is on developing scalable, integrated, cost-effective software tools and systems that embody and automate practices and processes supported by theory and research. Those theories, practices and processes must first be articulated precisely enough to support their automation. The articulated processes can exist first as guidelines, prescriptions, and decision aids, which can be turned into functional specifications for tools and systems. Once tools and systems are built, studies can be done to validate the theories, practices and the processes they embody.

The goal of this roadmap is to provide substantive guidance on methods to integrate state-of-the-art learning strategies into current medical simulation systems, identify unmet needs of both users and educators, and establish concrete goals for developing a new and more effective generation of simulators for learning, assessment and certification in medicine

2 A Brief History of Simulation-Based Training in Medicine

From the early 1990's, a number of parallel, non-convergent events led several medical domains to recognize that increases in computing power could lead to higher fidelity simulations for teaching cognitive, perceptual and motor tasks. Much of the emphasis was placed on procedural tasks and the technical skills required for performance of routine tasks in medicine.

The earliest plastic manikins incorporated computerized systems to control mechanitronic features such breathing, heart sounds and control of airway responses to user manipulations. There were seminal publications by authors such as David Gaba and DeAnda (1988) and simulation meetings held annually in Rochester, NY and elsewhere that were largely directly towards residency training in anesthesiology.

At the same time, funding from federal sources such as the Advanced Biomedical Technology program (Defense Advanced Research Projects Agency), led by Dr. Richard Satava and others, and the Visible Human Project (National Library of Medicine) were fueling developers for building virtual reality (VR), computer graphics and haptic devices (hardware providing the user with force and tactile feedback). Small companies were formed which eventually led to a plethora of prototypes (Higgins, et al., 1997a), most of which were never commercialized or produced viable products. On the software side, interactive media developers were considering relatively more simple software applications that could be used to train cognitive skills. Many of these ending up being more like 'electronic textbooks' than providing a paradigm shift in how the web and other media could transform the learning process in medicine.

The enthusiasm of the 1990's led many technology developers to believe that the culture of medical teaching could be transformed overnight into embracing the first generation of medical simulation products. Optimistic forecasts predicted rapid adoption of medical simulation, even using them as platforms for certification:

“When all the factors are put together, following the historical path taken by flight simulation, it is possible to forecast that surgical simulators will provide sufficient capability to serve as testing and certification instruments by the years 2005 – 2010” (Higgins, 1997b).

However, it soon became apparent that both culture and technology would have to evolve before these solutions would become part of normal medical education. Often, the products purchased by early adopters were left dormant sitting in the back of classrooms, especially since both content and reliability were often inadequate in the early product offerings.

The Role of Surgeons in the Establishment of Simulation Training

Some of the earliest advocates for simulation-based training in medicine have been surgeons and interventionists (Higgins, et al. 1997b; Higgins, 2002) because their operative procedures often involve manipulation of tools at-a-distance while watching interventions on a monitor or procedures at-a-distance. These domains have been enhanced by the advent of minimally-invasive surgery for replacing most general surgical procedures and the increasing use of interventional radiology, neuroradiology and cardiology methods. These settings lend themselves naturally to simulators where the operator is already operating at a distance following image-based guidance, with a configuration that typically includes tools that provide force feedback (haptics) to the user, manipulation of deformable computer graphics models of human anatomy, plus the ability for the simulator to objectively measure performance and provide feedback.

As advocates of simulation-based training, surgeons and interventionists may be well-positioned to lead a paradigm shift in simulation-based educational tools. The position of the surgeon has not changed much over the past 100 years – progression through a career path has relied on cognitive testing, peer review and mentoring for certification and procedural accreditation. Although technical abilities were tested in the early twentieth century (e.g., knot-tying), these were largely abandoned with the emergence of modern medical practice.

Now that computers, whose applications have been validated for analysis of task performance (Schraagen, 2000; Issenberg, et al., 2005) technical and cognitive skills can be objectively measured to an appropriate proficiency level, and the individual surgeon or other medical practitioner can be trained and assessed. Remediation and feedback are a critical component of this learning approach. The question is: Will this generation of senior surgeons and other master clinicians accept this paradigm shift? It will require them to embrace technology that they might not feel comfortable with, or changing long-standing curriculum in which they have invested much development.

Metrics and Proficiency

Validated metrics and measures are critical components of any education and training program. One of the biggest contributions of surgical simulation has been a well-laced focus on the same attributes for verification and validation that have worked with flight and defense simulation. In a landmark meeting in 2001, called “Metrics for Objective Measure of Surgical Skills”, a group of clinicians and others gathered, led by Dr. Richard Satava, to better understand what metrics are and how they can be applied in surgical skills training and assessment.

Among the many valuable results from the conference was: (1) a description of the different kinds of validity, based on the work of Raber, used in testing, and (2) a taxonomy of skills used in surgery and intervention. Table 1 summarizes the definitions of validity and reliability. Table 2 provides a summary of the taxonomy of abilities, skills, tasks and procedures.

Table 1. Definitions of Validity and Reliability

Term	Definition
Face validity	“... experts review the tests to see if they seem appropriate ‘on their face value’. EXAMPLE The chosen tasks resemble those that are performed during a surgical task
Content validity	experts perform “... a detailed examination of the contents of the tests . . . to determine if they are appropriate . . . and situation specific ...”. EXAMPLE: The tasks for measuring psychomotor skills are actually measuring those skills and not anatomic knowledge
Construct validity	“... the determination of the degree to which the test captures the hypothetical quality it was designed to measure ...”. EXAMPLE: The tasks were designed to test the level of a skill, therefore an expert should perform better than a student.
Concurrent validity	“... the relationship of the new test scores ... (and those) whose performance has been evaluated in actual working conditions ...”. EXAMPLE.: The scores on the test corresponds to scores on the current similar or “gold standard” tests.
Predictive validity	determining the extent to which the scores on a test are predictive of actual performance ...”. EXAMPLE: Those who do very well on the tests will do very well in the operating room.
Inter-rater reliability	“... determining the extent to which two different evaluators (raters) score the same test ...”. EXAMPLE: Two surgeons evaluate a student performing dissection of the gallbladder and both agree on the same errors, time, etc scores
Test-retest reliability	“... reliability of a test by administering it two (or more) times to the same persons and obtaining a (correlation) between the scores on each testing ...”. EXAMPLE: Students are tested twice on the same test and get equivalent scores each time.

•
•
•
•
•
•

Table 2. A Taxonomy of abilities, skills, tasks and procedures

ABILITIES		TASKS
Psycho motor Tracking Pick and place Translation Aiming (Targeting) Precision Visio-spatial Card rotation Cube comparison Map planning Depth Perception PicSOr Haptic		Anastomosis Bowel Vascular Laparoscopic Excise Superficial lesion Deep lesion (e.g. Breast) Closure (especially wound Tissue extraction Exploration (probing) Camera navigation Needle Insertion Aspiration Injection
SKILLS		
Instrument Handling Bimanual dexterity Transfer traversal Peg Board Navigation Ligation Foam Trotters Bowel (ex vivo) Vascular Tendon Laparoscopic Suturing Knot tying Open Extra-corporeal Incision Exploration (both visual and haptic) Cannulation Tissue handling Cutting Blunt Dissection Clamping (clip application) Hemorrhage control (simple) Plaster application		Pericardiocentesis Starting an IV Debridement (dissection) Morcellation Energy use (diathermy, scarifying) Stents Implant (prosthesis, mesh, etc) Hemorrhage Control Mesh Placement Evacuation Lymph node dissection Organ entrapment (eg Lap Bag)
		PROCEDURES
		Laparoscopic cholecystectomy Tracheostomy Pecutaneous Open Chest tube insertion Diagnostic peritoneal lavage Vein patech Breast biopsy (to be developed) Node dissection Ultrasonic diagnosis Endoscopic* Sinusoscopy Colonoscopy Bronchoscopy Arthroscopy Image guided Coronary stent

While this conference made good progress in defining the different kinds of skills and tasks that needed to be objectified for surgery and intervention, such work has rarely been undertaken for either academic prototype simulators or commercial products using approaches as such as cognitive task analysis (Seymour, 2002).

Gallagher and collaborators (Gallagher, et al., 2005), developed an approach to measure proficiency on an interventional neuroradiology simulator for carotid stent placement. This work led the U.S. Food and Drug Administration to require that a given proficiency be reached on the VIST simulator (Mentice, Inc.) before that individual could order the product in question. Subsequently the FDA has established its own internal board to see what other devices must pass such scrutiny before use by an operator. These actions may herald a whole new era where simulators are used for testing before, or instead of, human testing and application in a clinical setting.

Meta-Analysis Findings in Simulation-based Training in Medicine

Issenberg and colleagues at the University of Miami performed a literature review and attempted a meta-analysis of the literature on medical simulation since 1969 (Issenberg, et al., 2005). Similar meta-analyses of other domains, such as defense simulation and flight training reveals a large body of published literature in which training value of simulation has generally been well established with skill transference demonstrable from the simulator to the real world for a diverse array of tasks (Higgins, et al., 1997b; Higgins 2002). In contrast, Issenberg and colleagues found a relatively weak limited body of published literature on the efficacy of medical simulation for learning or debriefing, indicating that additional research is needed.

The Issenberg meta-analysis did, however, identify key features and uses of the medical simulations as educational intervention shared by many of the 109 studies reviewed. The findings, summarized in Table 3, reflect what is known from both the learning sciences and studies of simulation-based training in the military.

•
•
•
•
•
•

Table 3. Features and Uses of Medical Simulation as Medical Education Interventions

<p style="text-align: center;">Features and Uses of Medical Simulations As Medical Education Interventions Source: Issenberg, et al., 2005</p> <p>(1) Feedback. Meaningful assessment of performance was the single most important feature of the simulation systems studied. The studies indicated that feedback improved initial learning outcomes and also improves retention.</p> <p>(2) Repetitive practice. Nearly 40% of the journal articles reviewed for the meta-analysis cited the opportunity for repetitive practice with clinical problems and devices as a key feature of medical simulators. The use of medical simulators to practice skills was shown to improve skill acquisition and maintenance.</p> <p>(3) Curriculum integration. Simulation-based education was found to be most effective when it was integrated into the curriculum rather than used as optional exercises. 25% of the studies reported that simulation should be an integral part of the student's training program, rather than viewed as an extra-ordinary activity.</p> <p>(4) Range of difficulty level. It is important to provide training across a wide range of difficulty levels. 14% of the papers reviewed in the meta-analysis cited the need to support learners' through the learning curve – from basic skills levels through progression to mastery.</p> <p>(5) Multiple learning strategies. 10% of the studies identified the importance of designing medical simulations so that they can be adapted to support multiple learning strategies from instructor-centered lectures or instructor led group or individual exercises to independent learning.</p> <p>(6) Capture clinical variation. Medical simulations that can present a wide variety of patient problems or conditions increase the number and variety problems and situations that learners are exposed to. Simulations that can sample from patient demographics, pathologies and responses to treatment can present students with a rich variety of clinical variations.</p> <p>(7) Controlled environment. Nearly 10% of the studies cited the importance of providing a controlled clinical environment in which the learner can make and correct patient care errors without adverse consequences, and allow instructors to focus on the student, rather than the patient.</p> <p>(8) Individualized learning. Medical simulations make it possible to tailor learning experiences to meet individual learning needs. Nearly 10 % of the papers reviewed identified individualized learning as a key benefit of simulation-based learning systems. Complex clinical tasks can be broken down into their component parts so that students can progress to mastery in sequence at variable rates, yet all have the opportunity to achieve mastery.</p> <p>(9) Defined outcomes or benchmarks. Well-articulated metrics that are understood and accepted by the instructor and learner are critical. The defined outcomes must be observable and measurable in order to be meaningful. This feature was identified by 6% of the reviewed journal articles.</p> <p>(10) Simulator validity. The degree of realism or fidelity the simulator provides as an approximation to complex clinical situations, principles and tasks was identified as an important feature for training visiospatial perceptual skills and responses to critical incidents.</p>
--

3 Relevant Findings from Flight and Defense Simulation Training

Simulation has been successfully used by the military and for flight training for many years. At one end of this spectrum are flight simulators that cost millions of dollars and are used for selection and training of pilots. At the other end are part-task trainers, such as mannequins used to train medical personnel for medical emergencies specific to combat. A recent meta-analysis of flight and defense simulation training revealed robust studies demonstrating the efficacy of simulation. The study examined over total of 140 publications, including technical reports, scientific papers, and case studies, spanning more than fifty years. Special emphasis was placed on reviews of the literature that summarized transfer effects of simulation training. Other publications included studies of cost efficacy, technical description of simulators and features, and literature reviews.

The analysis identified and examined four key factors in the military and flight simulation literature, which are considered critical for the successful development and implementation of medical simulation-based training in the military. These factors include:

- Training Efficacy – The degree to which the skills trained in the simulation environment transfer to real world skills.
- Performance Assessment – Functions embedded into the simulator, which can track and measure the performance of the simulator operator.
- Fidelity – The extent to which the simulator reproduces the physical characteristics of the real world procedure, equipment or skill being simulated.
- Part-Task Training – Selective focus on the training of specific critical skills deconstructed from larger tasks.

Training Efficacy

A commonly used measure of simulation-based training efficacy is “transfer”, that is, how much student performance can be transferred from the simulator to actual, real world procedures. This can be expressed as a transfer effectiveness ratio (TER), which has been reported for a number of cases in the literature.

There are several ways in which simulation-training efficacy can be measured. Using direct simulator versus traditional method testing, the performance of a student trained only on a simulator can be compared to a control student tested using traditional mentoring procedures. Within-simulator performance evaluation is another common approach used for testing the training effectiveness of simulators. These data are then evaluated using statistical tests designed specifically to compare

•
•
•
•
•

simulation performance with actual performance in the real world. These data can then be used to determine transference rates.

The literature shows that, in general, the transference rate for a flight simulator is 0.48 - that is, 60 minutes using a flight simulator is worth 30 minutes flying an airplane in terms of training efficacy (Orlansky et al, 1994).

Many investigators have found simulation to be an extremely effective training instrument (Advisory Group, 1980; Carreta and Dunlap, 1998; Hays and Singer, 1989; Hayes et al, 1992a; Hays et al, 1992b; Jacobs et al, 1990; Knerr et al, 1986; Moroney and Moroney, 1998; Orlansky and String, 1979; Orlansky et al, 1994). These include studies in which a meta-analysis of the simulation literature has been performed to provide a more sensitive measure of the benefits of simulation training, using “field effects analysis” and other statistical methods. Although there is some controversy surrounding the validity of data pooled in such a manner, the results show that, in general, simulators provide an extremely valuable training effect comparable to training using actual equipment and real world procedures.

Performance Assessment

There are numerous studies that show performance assessment is critical to effective simulator training (Benton et al, 92; Caro and Isley, 1966; Caro et al, 1984; Connolly et al, 1989; Copenhaver et al, 1996; Dohma, 1995; Guckenberger et al, 1993; Hettinger et al, 1994, 1995; Jorna et al, 1992; Jacobs et al, 1990; Marcus and Curran, 1988; McCaulley and Cotton, 1982; Orlansky et al, 1997; Roscoe and Williges, 1980; Sterling, 1993a,b; Spears, 1983; Thomas et al, 1990; Westra et al, 1981; Westra et al, 1988). Performance measures may include simple functions such as listing the order in which a user activates a sequence of switches, or may involve sophisticated, computer-based systems that can evaluate users in a complex, distributed virtual environment. Metrics may include measures such as timing, accuracy, tissue damage, instrument handling, applied force, cognitive decision-making and others.

In military simulation, several factors have been found to be important for implementation of successful performance measurement systems in simulation. First, users benefit from being able to compare their performance against other users of the simulator, for example, other trainees or expert users. Second, performance measures on the simulator must relate to performance on the real world procedure being simulated. Finally, users must be motivated to perform well on the simulator. Thus, successful performance on the simulator should be tied to successful completion of the training requirement. In addition, features that provide real-time performance feedback to the user have been shown to enhance training.

Fidelity

Fidelity is the degree to which the simulator reproduces the actual, real-world procedure being simulated. Hays and Singer (1989) defined fidelity as:

“...the degree of similarity between the training situation and the operational situation which is simulated.”

The bulk of data show that physical (also known as ‘objective’) fidelity is not a requirement for successful simulation-based training (Advisory Group, 1980; Benton et al, 1992; Boldovici, 1987; Carretta and Dunlop, 1988; Copenhaver et al, 1996; Cyrus, 1978; Dixon and Curry, 1990; Dixon et al, 1990; Dohma, 1995; Durall et al, 1978; Edwards, 1986; Hays and Singer, 1989; Hays et al, 1992a,b; Knerr et al, 1986; Lees and Bussolari, 1989; Lintern, 1980; Lintern et al, 1987; Lintern et al, 1989; Lintern et al, 1990; Lintern et al, 1997; Loesch and Waddell, 1979; Martin, 1981; Martin and Waag, 1978a,b; Martin and Cataneo, 1982; McDaniel et al, 1983; Micheli, 1972; Montemerlo, 1977; Moroney and Moroney, 1998; Nataupsky et al, 1979; Orlansky et al, 1997; Pfeiffer, 1983; Rankin et al, 1984; Ryan et al, 1978; Voss et al, 1970; Waag, 1980; Westra et al, 1981; Westra et al, 1982; Westra et al, 1988; Westra, 1982).

Interestingly, several studies showed that students trained using low fidelity simulation can perform as well or better than students trained using high fidelity simulation:

- Caro (1988) showed that for novice training, simple wooden mockups were as effective as sophisticated cockpit simulators for training.
- Warren and Riccio (1985) showed that providing irrelevant stimuli in the context of a higher fidelity simulation actually made task learning more difficult as the novice trainee has to learn to ignore these stimuli.
- Kass, Herscheler and Companion (1991) showed that students trained in a “reduced stimulus environment” that presented only task-relevant cues performed better in a realistic battle field test than those who were trained in the battle field test condition.
- Lintern, Roscoe and Sivier (1990) showed that naive students trained without crosswinds in a simulated landing task performed better than students trained with crosswinds in landings that have crosswinds.
- Lintern and colleagues (Lintern et al, 1990b; Lintern and Garrison, 1992; Lintern et al, 1997) found that pictorial displays were more effective than symbolic displays in training landing skills, but increases in scene fidelity either had no effect on performance or actually reduced performance in some cases by distracting the trainee.

Although the overwhelming preponderance of data from military, flight and medical simulation show that simulators do not have to exhibit high fidelity to be useful

•
•
•
•
•

training instruments, they do have to have face validity for the end-user. Face validity is the degree to which the simulator appears 'genuine' and is adopted by the end-user. Often simulators will be used enthusiastically by trainees if they are endorsed by known content experts, and/or have a demonstrable ability to improve skill, even if the simulator appears unrealistic.

A classic case about simulation fidelity was the controversy about the need for motion platforms to realistically reproduce aircraft motion in flight simulation training. These motion platforms are very expensive, but are enthusiastically embraced by users as providing a much more realistic training experience than can be provided by static flight simulators. However, as reviewed in Moroney and Moroney (1998), the majority of the data show that, for most tasks that have been studied, motion platforms do not provide any additional instructional advantage over static systems. Boldovici (1992) interviewed 24 experts in the field and came to 11 conclusions about the need for motion platforms. He found, among other results, that greater transfer-of-training can be achieved by less expensive means than using motion platforms. Therefore, if cost is a requirement, motion platforms will never demonstrate an advantage. User's and buyer's acceptance is not an appropriate reason for the use of motion platforms.

The point is that the emphasis in simulator design and development must be focused on an accurate definition of the skills to be trained by the simulator, and not creation of the technically most realistic trainer possible. If the skills to be trained are adequately addressed, then low fidelity simulators may perform adequately, and the degree of fidelity required can be evaluated using the simulator.

Several authors have suggested that simulator fidelity be matched to the stage of learning: cognitive (initial), associate, and autonomous. Low fidelity simulators have been proven to be effective for initial training and sustainment training, whereas higher fidelity trainers may be more appropriate for autonomous learning.

Part-Task Training

The bulk of modern simulation training data show that part-task training is more effective for training difficult and "high performance" skills than is whole task training (Adams et al, 1962; Advisory Group, 1980; Aukes and Simon, 1957; Bailey et al, 1980; Gray, 1979; Kilion et al, 1987; Knerr et al, 1986; Mattoon, 1994; Naylor, 1962; Orlansky et al, 1994; Sheppard, 1985; Wightman, 1985; Wightman and Sistrunk, 1995; Wray, 1987).

Knerr et al (1986) reviewed the literature on flight training with regard to the use of simple, low fidelity trainers such as the Cockpit Procedures Trainer (CPT) and their utility for initial training and sustainment. The CPT is a part task trainer that is simple, low fidelity and inexpensive, and represents an early model for part task trainers that can be effective in terms of training value and cost. The results suggest that partitioning a large, complex task into complete, coherent parts does not disrupt learning and subsequent performance of the parts. However, it was felt that students needed to train a small amount on the entire procedure to learn time-sharing between

individual parts. The study concluded that:

- Part-task training of a skill that received very little practice in flight can be highly cost effective.
- CPTs (simple, low fidelity part-task trainers) have some value for *ab initio* training as long as student pilots have some opportunity to practice the whole task so that can acquire time-sharing skills.
- CPTs (simple, low fidelity part-task trainers) are very effective at sustaining procedural skills that are otherwise susceptible to forgetting over periods of no practice.
- CPTs (simple, low fidelity part-task trainers) may be effective for transition training of experienced pilots on the procedural aspects of new aircraft.

From these and other data, it is clear that part-task trainers can be used: 1) to train on complex procedures that require extensive practice to achieve proficiency, where critical steps (tasks) require “high performance”; 2) for sustaining training of procedural tasks, and to provide initial training on new procedures and tasks.

One other general finding is that part-task trainers do not require high fidelity, in part, because only a portion of the entire task needs to be simulated, and the emphasis can be placed on training a specific, highly critical skill, not on reproducing the entire procedure from start to finish.

4 Key Concepts from Learning Science

The science of learning should drive the specification of particular features of the learning environment that are appropriate to meet the learning objectives. For example, if research determines that training complex decision making skills requires extensive hands-on practice (a pedagogical conclusion) then developing a simulation to allow ample practice would be justified. New technology brings new educational opportunities, and these require changing existing curriculum. The new and *different* educational content should benefit from the *difference* in the technology, such that the curriculum has added value beyond the same material presented with previous technology. Simulator technology brings real time interactivity and "infinite perspective", which invites learning through discovery. Therefore, added educational value can be attained by creating highly interactive 3D models that provide an "infinite number of perspectives" (i.e. the anatomical structures can be seen from a limitless number of angles, from outside or inside, etc) and invite the student to learn by exploring and interacting with the anatomy. Compared to books and anatomical dissections, simulators bring the promise of individualized learning on patient models, so the content using simulation should emphasize this form of active interaction, rather than simply as an addendum to series of slides shown to a passive audience.

This section summarizes several key concepts from learning science and is intended to highlight research that show particular promise in optimizing simulation-based learning system. The goal of R&D addressed in Section 5 of this roadmap is to employ these key set of principals that have been demonstrated to enhance learning into next-generation medical simulation systems. This section is derived from a white paper, *The VMAS Educational Framework developed as part of the work of the Validation Methodology for Medical Simulation Training (VMAS) Committee for the Telemedicine Advanced Technology Research Center (TATRC) of the US Army Medical and Materiel Command* (Howell and Higgins ,2004), describing an educational framework for training combat medics, physicians and others to increase the readiness of medical personnel in the military.

Problem-centered Learning

One way to help students learn about conditions of applicability is to design problems that help students learn when, where, and why to use the knowledge they are learning. Proponents of using problems as a vehicle to contextualize learning suggest that transfer of learning will be better than instruction that presents content out of context. Using problems to anchor learning bridges the gap between general and specific knowledge since the general knowledge is learned in the context of specific applications. Active engagement with new knowledge and skills is an essential prerequisite to learning. By starting a lesson with a problem the engagement process begins right from the start. Starting with a problem makes learning a much more inductive experience, especially when the learner has multiple options to build the knowledge base needed to solve the problem. Sterling (1996) emphasizes the importance of case studies for simulation-based training.

Learning is promoted if the instruction provides a structure that the learner can use to build the required organizational schema for the new knowledge. Merrill (2002) recommends starting with easier problems and moving to more difficult ones and providing more support in the form of hints and demonstrations available in the beginning and removing support as learning proceeds. Andre (1997) discusses the role of advance organizers in providing structure for later learning. Mayer (1975) indicates that providing learners with a conceptual model can facilitate the acquisition of problem-solving. Clark and Blake (1997) recommend presenting dynamic schema and analog models to promote far transfer.

The challenge in instruction is to provide learning environments that manage the limited processing capability in working memory so that new information gets encoded into long-term memory in a way that it can be effectively retrieved or transferred later. Experts in a subject domain typically organize factual and procedural knowledge into schemas that support pattern recognition and the rapid retrieval and application of knowledge (Chase and Simon, 1973; Chi et al., 1981). Experts' abilities to solve problems depend strongly on a rich body of knowledge about subject matter that support thinking about alternatives that are not readily available if one only memorizes facts. (Bransford and Stein, 1993). Experts have not only acquired knowledge, they are also good at retrieving the knowledge that is relevant to a particular – conditionalized knowledge is knowledge that includes a specification of the contexts in which it is useful (Glaser, 1992).

Critical Skills Focus

It is critical to understand expert-novice differences and ensure the curriculum addresses those tasks that make substantial impact to critical job performance and that require demonstrations and practice to learn. Experts can rarely articulate the mental models that are the source of their expertise. They have so much tacit knowledge stored in long-term memory that it is difficult for them to explain it verbally. For example, detailed analysis estimated that chess masters have about 50,000 play patterns stored in their long-term memories, patterns routinely used as the basis for game strategies (Simon and Gilmarin, 1973). The military developed Cognitive Task Analysis (CTA) to facilitate rapid and effective acquisition of expertise by enlisted personnel in complex cognitive-technical skills (i.e. fighter pilot training, complex electronics trouble-shooting). CTA uses a structured interview and analysis process in which experts are asked to solve authentic job problems and at the same time to verbalize their problem-solving thoughts (Jonassen, Tessmer, and Hannum, 1999).

CTA studies can reveal performance differences between experts, intermediate-level learners and novices. Past studies have revealed differences in content and structure related declarative knowledge, knowledge schemes, pattern recognition, etc., corresponding to differences predicted from study of cognitive psychology and expertise.



Expert-novice differences were frequently in categories of assumed pre-requisite knowledge or learned solely through procedural knowledge.

Varied and Contrasting Examples

Merrill, Tennyson and Posey (1992) indicate that the use of well-chosen contrasting cases can help learners learn the conditions under which new knowledge is applicable. When teaching problem-solving or decision-making tasks, several examples that look different on the surface but that illustrate the same guidelines help to maximize transfer. The goal of instructional methods is to build mental models in long-term memory that will transfer effectively to working memory after training. Training can build specific mental models that apply only to limited situations or more flexible mental models that transfer to various situations. When training tasks that involve decision-making and problem-solving, a more flexible mental model gives better performance since it transfers to various diverse situations.

A number of studies by Sweller, van Merriënboer, and Paas (1998) have shown that training time can be reduced and learning improved when worked examples are substituted for some practice problems. Learning load is reduced and learning is made more efficient by using worked examples to build new mental models rather than spending working memory resources to solve problems, van Merriënboer defined an approach that incorporates both near and far transfer tasks. During the analysis phase, a top-down job and task analysis is used to define far transfer tasks (called nonrecurrent tasks) and near transfer tasks (called recurrent tasks), as well as the supporting knowledge for both including concepts, facts, mental models, and problem-solving approaches (called systematic approaches to problem solving). The job functions are then sequenced from simpler versions of whole authentic tasks to more complex versions. The problems presented are diverse in surface structure to help build more transferrable mental models.

Clark and Blake (1997) show that far transfer is promoted when the structural features are carefully identified and explicitly mapped for the learner; such guidance focuses the learner's attention on relevant information in the task. As the instruction progresses this information focusing should be faded and learners expected to attend to and focus their own attention on the relevant aspects of the information (Andre, 1997).

Demonstration

Several studies suggest that effective instruction should provide an opportunity for learners to demonstrate their newly acquired skills (Gardner, 1999); Perkins and Unger, 1999) and (Schwartz, et al, 1999). Instruction is far more effective when the information is demonstrated via specific situations or cases. Jonassen (1999) recommends demonstration of each of the activities involved in a performance by a skilled (but not expert) performer. He identifies two types of modeling: behavioral modeling which describes how to perform the activities identified and cognitive modeling which articulates the reasoning that learners should use while engaged in the activity.

Practice opportunities

Merrill (1994) cites research that shows that presenting examples in addition to practice promotes better learning than practice alone. Learning is most effective when people engage in deliberate practice that includes active monitoring of one's learning experiences (Ericsson et al, 1999). Research shows that adding practice to information and examples increases learning. Gardner (1999) and Perkins and Unger (1999) both emphasize the necessity of many opportunities for performance.

Reflection

The process of reflecting on and directing one's own thinking is one of the hallmarks of expertise. The Vanderbilt Cognition and Technology Group (Schwartz, et al, 1999) found that reflection is key to the integration of new knowledge and skills. The ability to recognize the limits of one's current knowledge, then take steps to remedy the situation is critical. Research shows that training learners to self-explain examples consistently improves learning outcomes. When learners actively study and encode the example new mental models are actively constructed and learning is maximized. Learners' thinking should be made visible through discussions, text or tests and feedback must be provided. The learning environment should incorporate techniques that require learners to self-explain examples to promote deep processing and maximum learning from examples.

Feedback

Feedback is most valuable when students have the opportunity to use it to revise their thinking as they are working on a task (Barron et al, 1998; Vye et al, 1998). Learners acquire a skill much more rapidly if they receive feedback about the correctness of what they have done. If incorrect they need to know the nature of the mistake. Timely feedback is critical so that the learner's practice of a skill and its subsequent acquisition will be effective and efficient. Feedback should occur continuously, but not intrusively, as a part of instruction. Technology is providing new learning tools that can be used to monitor actions, intervene with hints and feedback, ask questions to elicit learner understanding, and direct learners to summon instructors when the learners need additional help (Genscope, Hickey, Kindfield and Horwitz, 1999).

Assessment

Assessments function within a large system of curriculum, instruction, and assessment. Changing one of these elements and not the others runs the risk of producing an

incoherent system. All of the elements and how they interrelate must be considered together. Every educational assessment should be based on a set of foundations: 1) every assessment is grounded in a theory about how people learn, what they know, and how knowledge and understanding progress over time; 2) each assessment embodies certain assumptions about which kinds of observations, or tasks, are most likely to elicit demonstrations of important knowledge and skills from students; and 3) every assessment is premised on certain assumptions about how best to interpret the evidence from the observations to draw meaningful inferences about what students know and can do.

Most assessments are “static”; they provide snapshots of achievement at particular points in time, but they do not capture the progression of students’ conceptual understanding over time, which is at the heart of learning? This limitation exists largely because most current modes of assessment lack an underlying theoretical framework of how student understanding in a content domain develops over the course of instruction, and predominant measurement methods are not designed to capture such growth. Assessments should include learners’ organization of knowledge, problem representations, use of strategies, self-motivating skills, and individual contributions to group problem solving.

A general paradigm for conducting experimental evaluations of simulator training effectiveness for medical applications is based on the work of Pugh, Hettinger, Higgins and others (Pugh et al, 2001a,b; Hettinger et al, 1995; Higgins et al, 1997; Lathan et al, 2001). It uses the Transfer-of-Training paradigm that has been successfully applied to simulator training evaluations in aviation and other domains (Moroney and Moroney, 1998; Orlansky et al, 1994; Champion and Higgins, 2000). This work is being coordinated with the VMAS (Validation, Metrics, Assessment for Simulation) Steering Committee.

Skills Refreshment

There are “learning curves” for the performance of medical procedures, but these may vary between individuals. In a laparoscopic procedure such as cholecystectomy, the steepest part of the learning curve has been empirically demonstrated to be the first 10 cases the surgeon performs, but can continue up until the first 50 cases have been completed (The Southern Surgeons’ Club 1995); for GI endoscopy, the learning curve has been estimated to be as many as 300 procedures (Cass 1999). The urology learning curve for procedures such as cystoscopy has been estimated to be 25-100 procedures (Shah and Darzi, 2002). There has been considerable debate about the need for increased procedural volume to reduce error in the performance of procedures such as percutaneous transluminal coronary angioplasty (PTCA) (Hannan et al, 1997; Ritchie et al, 1993; Ryan et al, 1996; Shook et al, 1996). Initial studies showing that adverse outcomes were significantly higher in low-volume centers (Ritchie et al, 1993; Jollis et al, 1994), raised public concern (Squires, 1996), but did not adequately explore the issue of whether low-volume centers may have more complications because they treat high-risk patients (Ryan, 1995). Two important findings of the work of Kimmel et al (1995), as emphasized by Ryan (1995), are that laboratory volume was linearly and inversely associated with major

complications, and the odds reduction from major complications becomes statistically significant when laboratory volume is more than 400 cases per year.

Professional organizations such as the American College of Cardiology and American Heart Association have recommended that physicians practice a minimum number of PTCA procedures per year as the primary operator in order to maintain competence in the procedure (Douglas et al, 1993; Ryan et al, 1990). In Maryland, civilian EMT medics are required to perform at least 6 needle sticks a year to maintain certification. For many practitioners, it may be difficult to meet the professional guidelines for maintenance of competence established because they do not treat enough patients, thus training on a simulator may take the place of performance on patients to help prevent skills decay. However, trainers or professional organizations that require trainees to ‘train’ on simulators without much systematic thought about what they are trying to achieve may not recognize individual variability. The underlying assumption seems to be that individuals who have performed the required number of procedures will be safe practitioners, but this ignores variability in individual learning rates. Setting a fixed number of procedures or number of training hours is a less than optimal approach to learning.

Reusable education and training materials

The development of education and training materials is time consuming. It involves contributions by many disciplines: domain experts, information technologists, computer scientists, cognitive scientists and education and training specialists. Some parts of the tools and content are common among different courses, some are unique. Considerable effort and cost can be saved if the tools and content are reusable. If designed with re-use in mind, the components and objects developed can be used by different instructors, teachers and learners for different purposes. Re-usable simulations, instructional content and learning tools can be combined to create new learning systems, can facilitate maintenance of learning systems by making it easier to keep material up-to-date, and allow customization of learning systems for specific learners’ needs (FAS, 2003). In addition, resources can be used in ways not previously considered, permitting creation of totally new teaching resources. Simulations and content should be structured in such a way that they can be adapted and reused.

Content material should include metadata to permit search and retrieval from digital repositories. Simulations should be designed to be interoperable so that larger systems can be built from component simulations. Current work in learning objects, interoperable simulations, software development communities and digital libraries should be used and expanded. Examples include the Department of Defense’s Advanced Distributed Learning (ADL) Initiative (ADL, 2005), which is designed to accelerate large-scale development of dynamic and cost-effective learning software and systems to meet the education and training needs of the Military Services through the development of a common technical framework for computer and net-based learning for the creation of

•
•
•
•
•



reusable learning content as "instructional objects." The Federation of American Scientist's Digital Human project is focused on building a community of researchers working in biomedical simulations to develop a framework to support interoperable software components and biological simulations (FAS Digital Human, 2005).

5 Research Challenges and Topics

Despite important successes in the use of simulation-based medical and training systems there are still a number of limitations in our knowledge of how best to take advantage of this new approach to learning that pose several important challenges for their routine use in learning environments. In addition, sophisticated computer software tools are needed— software ranging from simulations of biological processes to systems capable of creating learning content tailored to the individual and to specific learning objectives. Developing these software tools and systems will be like other software development efforts: difficult, labor-intensive, and expensive. Building these specialized tools is far beyond the capacity of most instructional designers. Tools to decrease the level of effort are desperately needed. Frameworks and guidelines for how to design and integrate simulation-based learning into medical curriculum are needed, as well as large-scale demonstrations that support evaluation of the approaches and software tools.

The major design objectives of the R&D needs that are described in this research agenda include:

- 1) Generation of scenarios and cases *on-the-fly* and tailored to address an individual instructional or programmatic need.
- 2) Case scenario authoring that hides the underlying technology and can be easily used by instructional designers, educators, and parishioners.
- 3) Intelligent design of training systems, including automatic selection of the most appropriate instructional techniques for each type of learner and training objective.
- 4) Creation of patient case/scenarios through aggregation of learning content from multiple sources, including automatic retrieval from clinical information, including “visual clues” drawn from examination of the appearance of the patient, as well as vital signs, brief medical history, lab results and reference diagnostic images.
- 5) Motivating learning systems that engage learners and increase practice of skills.
- 6) Enhanced learning and practice systems for complex decision-making in teams, including capabilities that permit team members to develop and execute strategies, adapt to changing events, and develop compatible mental models of teammates’ roles, the tasks, and the situations the team encounters.
- 7) Rich, meaningful profiles of learners that enable accurate analysis of mastery, feedback, coaching, and prescription of content during instruction.
- 8) Effective response to learner actions and queries by giving quality feedback/answers in a variety of forms including verbal, textual, and visual.

•
•
•
•
•

-
- 9) Certification of the trainee's performance that can be used to support decisions related to education and training needs, selection, promotion, and performance management in organizations.
 - 10) Immersive environments that reduce the need for manipulation of computer peripherals.

The research priorities described in this roadmap are derived as a subset of the research priorities identified in the Learning Federation's 2003 Learning Science and Technology R&D Roadmap www.fas.org/learningfederation. The Learning Science and Technology (LS&T) R&D Roadmap, produced by the Federation of American Scientists' Learning Federation Project, described a vision for next-generation learning systems, and outlined a national research plan to radically improve approaches to teaching and learning through information technology. The LS&T R&D Roadmap was produced over a two-year period with input and advice from over seventy researchers from industry, academia, and government through their participation in focused workshops, interviews, and preparation of technical plans. Comprised of a series of five component roadmaps, the Roadmap provides an assessment of R&D needs, identifies key research questions and outlines a chronology of R&D activities designed to spur innovation in technologies for education and training.

The Virtual Patient research road map is organized into *three* key research topics and the associated R&D tasks that were identified by the June 2005 workshop participants. These are:

- Automated generation of patient case scenarios
 - Frameworks for identifying the essential features of scenarios/cases that make them good learning tools.
 - Guidelines for linking scenario features with task types and learner characteristics, and
 - Automated scenario/case authoring tools
- Modeling Learners and Intelligent Tutoring
 - Common frameworks to define competency models
 - Automated tools collecting performance data and monitoring performance
 - Automated tools for translating cognitive task analysis data into diagnostic models
 - Guidelines and tools to optimize the introduction, format, timing and fading of assistance to the learner
 - Coaching strategies that dynamically adjust according to learner achievement
 - Automated processes for generating and presenting on-line feedback that is sensitive to the task and to learners
- Feedback and Evaluation
 - Guidelines for personalizing feedback based on the model of the learner
 - Interfaces for authoring feedback mechanisms
 - Standard data structures and transfer protocols to support combining and reporting assessment data, for the user and the learning systems

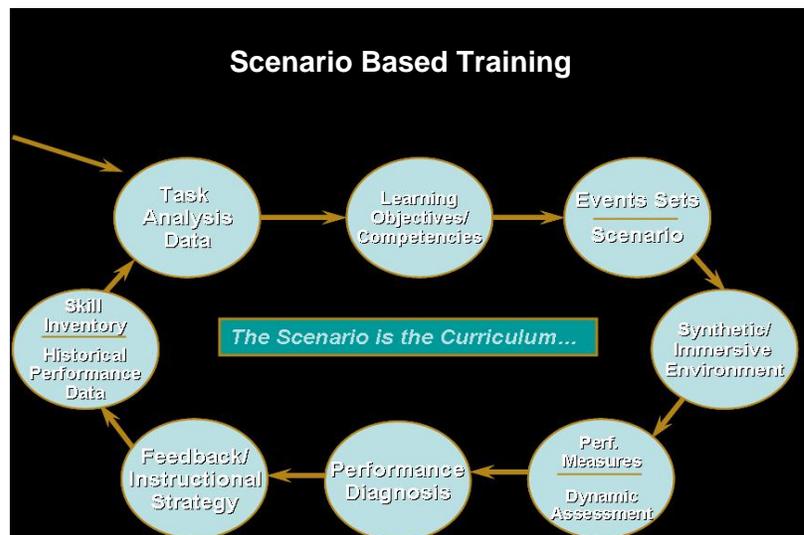
- Decision-aids for choosing different types of measurements and level of granularity based on context, budget and purpose

Collectively these research topics should facilitate effective simulator-based education and training, as described in the design goals. The following sections discuss each research topic and the associated R&D tasks.

5.1 Research Focus Topic#1: Automated Generation of Patient Case Scenarios in Medical Simulation

A central question in designing simulation-based education and training environments involves the design of scenarios and cases that can optimize the learning objectives required by the specific learner, serving as the benchmark for medical instruction. Working in a team training environment, Cannon-Bowers, Salas and colleagues (Cannon-Bowers and Salas, 1998; Salas and Cannon-Bowers, 1997; Fowlkes et al., 1998) developed a framework whereby specific trigger events are scripted into a scenario based on the learning objectives to be accomplished. These trigger events are designed to elicit desired behavior, to allow trainees to practice targeted skills and to provide an opportunity to measure performance and deliver specific feedback (Salas and Cannon-Bowers, 2000). To date, the event-based approach to training has been successfully demonstrated in several settings (Fowlkes et al., 1994; Johnston, Cannon-Bowers and Smith-Jentsch, 1995; Dwyer, et al., 1999). In a similar vein, Schank et al. (2000) advocate a strategy for developing goal-based scenarios or cases. This process includes guidance for developing learning goals, missions, cover, role, scenario operations, resources, and feedback. To date, tools to automate such processes have not been developed. Figure 4 describes the components of a typical training scenario.

Figure 4. Components of effective training scenarios. Source: Cannon- Bowers, 2005.



•
•
•
•
•
•

We identify three R&D priorities for generation of patient case scenarios in medical simulation:

- Frameworks for identifying the essential features of scenarios/cases that make them good learning tools.
- Guidelines for linking scenario features with task types and learner characteristics, and
- Automated scenario/case authoring tools

5.1.1 Research Task #1: Frameworks for identifying the essential features of scenarios/cases that make them good learning tools.

A program of research is needed that aims to elucidate and investigate in a systematic way the host of variables that will have an impact on instructional effectiveness for a particular educational or training goal. Efforts to develop and validate generalized instructional strategies or approaches are needed to guide instructional design for similar tasks, learning objectives and learners.

In general, such an approach would be concerned with what is being taught, who the learners are, which phase of instruction is of interest, how best to teach targeted material, the context in which learning will occur, and any practical considerations that limit what can be done. One way to organize a systematic program of this sort is to construct an overriding model or framework that lays out all of the pertinent variables and the manner in which they are related.

Longer term needs focus on development of models and vocabularies that can provide researchers a common framework in which to conceptualize their studies and make it easier to see how individual studies (i.e., the specific variables and context being tested) fit into the larger picture. In addition, a common framework will allow research results to be more effectively integrated across factors and gaps in understanding to be identified.

The following key research sub-tasks are needed:

(1) Identify the essential features of scenarios/cases that make them good learning tools. A central question in designing simulation-based practice environments involves the design of scenarios or cases as the backdrop for instruction. Efforts to develop and validate generalized instructional strategies or approaches are needed to guide instructional design for similar tasks, learning objectives and learners. Working in a team training environment, Cannon-Bowers, Salas and colleagues (Cannon-Bowers and Salas, 1998; Salas and Cannon-Bowers, 1997; Fowlkes et al., 1998) developed a whereby specific trigger events are scripted into a scenario based on the learning objectives to be accomplished. These trigger events are designed to elicit desired behavior, to allow trainees to practice targeted skills and to provide an opportunity to measure performance and deliver specific feedback (Salas and Cannon-Bowers, 2000). Schank et al. (2000) advocate a strategy for developing goal-based

scenarios or cases. This process includes guidance for developing learning goals, missions, cover, role, scenario operations, resources, and feedback. Current knowledge in how best to create effective simulation-based learning environments is not specific enough to provide robust guidelines for instructional designers. Research priorities include development of frameworks for identifying the essential features of scenarios/cases that make them good learning tools and guidelines for linking scenario features with task types and learner characteristics.

(2) Identify factors that contribute to learner motivation. Learning researchers are increasingly making a distinction between rote memorization of knowledge, and the more elusive goal of learning for understanding. According to Bransford, Brown and Cocking (1999), modern views of learning recognize that learning for understanding requires more than knowledge of a series of disconnected facts. Rather, these authors point to research that suggests that expert knowledge is “connected and organized around important concepts,” that is “conditionalized”, and that it supports the ability to transfer to other areas or domains.

Learner motivation has been linked to the meaningfulness of targeted material. In fact, many researchers converge on the conclusion that learning is enhanced when students are presented with relevant, meaningful learning goals and problems (e.g., see Bransford, Brown and Cocking, 1999; Clark and Wittrock, 2000; Perkins and Unger, 2000; Jonassen, 2000; CGTV, 2000). Also see Merrill (2003), Jonassen (2000) for further discussion. In addition, learning is enhanced when it occurs in a context that is meaningful to learners (Bransford, Brown and Cocking, 1999). Anchored or situated instruction is preferable because new learning can be more easily integrated into existing knowledge and mental models. Anchored instruction seeks to help learners understand the types of problems and opportunities that real experts confront and how they use their knowledge to solve those problems, and to help students integrate their knowledge by employing multiple perspectives on the same problem. Anchored (situated) learning environments allow learners to understand how new information is connected to what they already know.

(3) Provide guidelines regarding immersion and engagement requirements. Research associated with creating the appropriate degree of authenticity in the learning has been conducted and indicates that as with other learning environment features, which of these types of fidelity is important in learning depends on the nature of the learning objectives driving the instruction. As discussed in Section 3 several studies involving flight simulators have shown that students trained using low fidelity simulation can perform as well or better than students trained using high fidelity simulation. Jonassen (2000) discussed the notion of authenticity by pointing out that it does not necessarily mean that the instruction is developed around specific, real-world tasks. Rather, authenticity can best be thought of as the degree to which the learning environment causes learners to engage in cognitive processes that are similar to those in the real world (Honebein et al., 1993; Duffy and Savery, 1996; Petraglia,

•
•
•
•
•
•

1998). Moreover, authentic learning environments are engaging to learners, and provide them with challenging problems to solve (e.g., CTGV, 2000).

The underlying issue is related to the transfer of specific knowledge and skill to the actual operational or job environment. Specifically, if trainees are learning how to apply a particular skill, then the training (simulated) environment must respond in a manner that is similar to what would occur in the real world. Otherwise, the trainee will receive incorrect feedback and perhaps learn the wrong things. In this regard, Hays and Singer (1991) distinguish between physical fidelity (i.e., the degree to which physical features of the simulation are represented such as knobs and buttons) and cognitive fidelity (i.e., the degree to which the simulation faithfully represents conceptual aspects of the actual task). The concept of “selective fidelity” put forth by Andrews, Carroll and Bell, 1995) states that fidelity levels should be selected based on the learning needs of the learner. Learning outcomes are improved only if the correct cues and stimuli are presented to support the specific learning objective.

Dialogue and conversation are another important component of fidelity in virtual patients. Students and trainees need to be able to “converse” with virtual patients in order to identify patients’ problems and to learn how to communicate to patients their conditions, treatment options, and care. Verbal cues are a critical source of patient information. Dickerson, et al., 2006 studied the use of synthesized and recorded speech to better understand the tradeoffs in flexibility, fidelity and cost for use as virtual patients for medical education and training. Their study indicated that for lower level learning of communication skills, such as teaching students which questions to ask, synthesized speech (a low cost solution) was as effective and recorded speech. For teaching higher level communications skills such as how to ask the correct questions, a high level of expressive is required, thus indicating a more sophisticated dialogue system is required. Research in conversational dialogue is on-going in many areas of simulation and entertainment (Ruttkay and Pelachaud, 2004.) Key research needs identified here are focused on how to exploit the research, find ways to collaborate with on-going research, and approaches for integrating it effectively and establishing criterion for evaluation of its use in medical-based education and training.

Table 4 summarizes the key R&D topics for identifying the essential features of scenarios/cases that make them good learning tools.

Table 4. R&D topics to identify the essential features of scenarios/cases that make them good learning tools

Tasks	Milestones		
	3-years	5-years	10-years
Scenario/Case Design Features:	<p>Design of a pick and pull tool that is validated for the community</p> <p>Generally accepted definition of scenario/case for community to facilitate later adoption</p> <p>Standards and tools that enable visual cues drawn from the clinical examination of the appearance of the patient, as well as vital signs, brief medical history, laboratory results and ability to reference diagnostic images</p>	<p>Demonstrations that incorporate the essential features that improve learning outcomes</p> <p>Guidelines for optimal mix of features for various learning objectives</p>	<p>Community accepted frameworks for implementing features that improve learning</p> <p>Demonstration of increased effectiveness of instruction that incorporates features identified to improve learning, including transfer and retention</p>
Motivation: Identify factors that trigger motivational processes that increase likelihood that instruction will be successful	<p>Empirical studies that isolate which features, such as challenges, story, goal orientation, enhance motivation</p> <p>Results reflect learning increases by 10%</p>	<p>Empirical results linking motivational features to task type and learner characteristics</p> <p>Guidelines and mechanisms for assessing motivational features</p> <p>Results reflect learning increases by 25%</p>	<p>Validated modeling strategies for assessing motivation</p> <p>Demonstrated techniques to increase motivation across tasks and learners</p> <p>Demonstrated increase time/cost savings and learning ability: 25%</p>
Immersion/Engagement: Identify appropriate degree of authenticity to meet specific learning objectives	<p>Empirical results demonstrating the impact of immersion and engagement on learner motivation</p> <p>Empirical results defining immersion and engagement as viable psychological variables</p>	<p>Demonstration of psycho-metrically-sound techniques for assessing immersion and engagement</p> <p>Delineation of game features that foster immersion and engagement</p>	<p>Validated guidelines for assessing the degree of immersion and engagement</p> <p>Validated guidelines for increasing immersion and engagement</p>



5.1.2 Research Task #2: Guidelines for linking scenario features with task types and learner characteristics

Most modern theories of instruction converge on the conclusion that the attributes that learners bring to the instructional environment are important ingredients in the learning process. In fact, the notion that learners bring a unique set of knowledge, skills, preferences and experiences to a learning environment is captured by a popular approach known as learner-centered instruction (e.g., see CGTV, 2000; Bransford, Brown and Cocking, 1999 et al; NRC; Kanfer and McCombs, 2000; Clark and Wittrock, 2000). Essentially, proponents of this approach argue that characteristics of the learner must be taken into account in the design and delivery of instruction, and that an explicit attempt must be made to build on the strengths of the student. Variables that have been implicated in this regard include: prior knowledge, prior skill, prior experience, misconceptions, and interests.

Learning is enhanced with teachers/instructors pay attention to the knowledge and beliefs that learners bring to a task. Prior knowledge influences learning and research has shown that learners construct concepts from prior knowledge (Resnick, 1983; Glasserfeld, 1984) Prior knowledge affects how learners interpret instruction – research has shown that learning proceeds primarily from prior knowledge and only secondarily from presented materials. Learners’ misconception can result in the student learning something opposed to the educator’s intentions (Genter, 1983). Thus it is important that the design of instruction include approaches to make the learner’s thinking visible.

In learning systems, goals help to focus learners on the task and help them to select or construct strategies for goal accomplishment; hence, they serve to direct attention (Locke and Latham, 1990). Goal commitment (i.e., the degree to which the learner is committed to the learning goal) is a determinant of how much the goal affects performance (Locke et al., 1981). Therefore, efforts to understand goal setting and acceptance in TELS have the potential to increase learner achievement (Kanfer and McCombs, 2000). In addition, goal setting has been linked to self-regulatory processes (Schunk, 2001).Eccles, Wigfield and colleagues (e.g., Eccles, 1984; Eccles et al., 1993; Wigfield and Eccles, 1994) argue that subjective task values affect learner motivation. Several types of subjective task values are important: attainment value (i.e., the importance of doing well on a particular task); intrinsic value (i.e., enjoyment experienced by engaging in a task) and utility value (i.e., perceived usefulness of achieving a task). Similarly, many work-oriented theories of the construct “motivation to learn” are tied to Vroom’s (1964) theory of valence, instrumentality, expectancy (VIE). In particular, it has been shown that learners who see value in doing well in learning (in particular, when they believe that it will lead to some desire outcome) do perform better. These are important concepts that need to be considered when developing learning systems.

Table 5 describes the research tasks for linking scenario features with task types.

Table 5. Research topics for linking scenario features with task types.

Tasks	Milestones		
	3-years	5-years	10-years
Automated tools for task analysis: Link task demands to knowledge types and learning objectives	Specification of data types and collection methods Synthesis of existing techniques into initial design	Procedures for collecting data including methods to streamline knowledge elicitation and protocol analyses Procedures expanded to team tasks	Formative evaluation data on tool effectiveness Empirically-validated tool
Measurement tools: Develop assessment tools for learner characteristics	Synthesize literature regarding learner characteristics that are likely to have an impact on learning outcomes Prioritized list of candidate studies Preliminary assessment tools with data	Validated assessment tools Validated tools reflect 50% increase in captured information	Fully adaptable assessment tools
Ability/spatial ability: Determine how general, spatial, technological ability affects learning system design	Prototype of tools that can baseline user ability across 3 parameters Evaluate techniques for remediating 3 parameters when necessary	Refined tools for assessing ability online Prototype tools that can minimize potential limitations for participating in simulation-based	Open-standard software that can be embedded in 3 rd party simulations that diagnose and remediate across 3 parameters
Prior knowledge: Determine how preexisting knowledge and experience affect the system design	Empirical results of the relationship between prior knowledge and experience and learning system design	Refined tools for automatically assessing prior knowledge and experience prior to entering the learning system training/education	Automated processes for adjusting the learning environment in accordance with the learner's initial knowledge and experience
Goal orientation: Determine how goal orientation affects the learning system design	Empirical results of the relationship between goal orientation and aspects of simulation-based learning systems across 3 subject domains/skill classes	Develop guidelines within and across subject domains/skill classes	Embedded, open source tools to adjust curriculum/instruction based on goal orientation and content



5.1.3 Research Task #3: Automated Scenario/Case Authoring Tools

Case scenarios are expensive and resource-intensive to create. As result, learners are typically exposed to a limited number of cases. For computerized patient manikins this is less of a problem, as the major suppliers offer configurability, either in the form of open authoring capabilities or real-time adjustment of parameters as the patient’s condition runs its course. For other applications, authoring may take the form of choosing from a menu of options or programming macros. However, configurability still remains a problem, especially as these simulation systems may be localized to a variety of different settings, specialties and locales. The current state-of-the-practice is to use canned scenarios to provide the simulated environments. Ideally, we would like to be able to automatically generate scenarios that can be tailored to the user (individual or unit) educational or training needs.

Research shows that adding practice to information and examples increases learning. Gardner (1999) and Perkins and Unger (1999) both emphasize the necessity of many opportunities for performance. (Ericsson et al, 1999 demonstrated that learning is most effective when people engage in deliberate practice that includes active monitoring of one’s learning experiences. Merrill, Tennyson and Posey (1992) indicate that a necessary condition for effective concept instruction is was a range of divergent examples. The use of well-chosen contrasting cases can help learners learn the conditions under which new knowledge is applicable. The goal is to build mental models in long-term memory that will transfer effectively to working memory after training; a more flexible mental model gives better performance since it transfers to various diverse situations.

A key research challenge is to reduce the cost of creating scenarios, and thus increase the number of scenarios a student can experience and increase practice opportunities for students. The R&D can be described as three sub-tasks:

- Methods and tools to reduce the time and cost required to create virtual patient data for scenarios.
- Tools to make it easier to locate virtual patient data and to assemble data to meet specific learning objectives.
- Frameworks to enable interoperable virtual patient physiological models.

(1) Speed the creation of VP content, making it less costly. Virtual patients (VPs) are a key component of learning content for patient case scenarios. A VP is a set of data that describes an individual as a patient. The data can describe a real patient (a representation of a patient derived from a real patient’s clinical record), or a hypothetical patient created to address a particular topic or educational objective (Medbiquitous, 2005). For medical simulators, virtual patients are

computer-based simulations that simulate patient cases. VPs are expensive and resource-intensive to develop. As a result, efforts are underway to promote sharing, such as the American Association of Medical Colleges Virtual Patients Reference Center.¹ In addition, a number of individual and groups are engaged in activities to define data models for patient/case information that can be shared with applications, including the many different frameworks and vocabularies used for electronic patient records, such as MeSH, SNOMED, HL7, HEAL and MedBiquitous.

Many scenarios will require avatars, or virtual agents, that can either serve as the incarnation of the learner in the virtual environment or a virtual agent that interacts with the user as he/she moves through the environment. Within medical simulations avatars can be used as virtual patients (Hubal, et al., 2000) as virtual team members to substitute for training with real teammates, or as an instructor or coach, such as within an intelligent tutoring system (Grasser 2001; Rickel, 2001). Increasingly, researchers are using advanced technologies and modeling techniques to insert realistic human actors into simulations. Based on human performance and cognitive modeling techniques, this work has been on-going in the military for several years (Knerr et al., 2002). The benefit of this strategy is that it can heighten the authenticity of the learning experience by allowing trainees to practice higher-order skills with realistic actors who behave in an accurate, believable manner. These computer-generated actors can provide a low cost alternative to more traditional role-playing strategies by reducing the need for human actors. It can also allow team members to practice effectively, even when live teammates are unavailable. More work is needed to realize these potentials. Ideally, avatars should exhibit personality, emotion, self-motivations, adapt, and display variety in movement and responses. As with dialogue and conversational agents, research is ongoing. Similar to the key research needs identified for dialogue and conversational dialogue, the long-term goal is to leverage the progress being made in other fields so that generic avatars with these basic capabilities could then be customized for particular learning applications by programming specific skills, knowledge, or personalities to support their use in medical simulation education and training.

There has been some suggestion that patient case scenario authoring should employ a common, open source approach. This is somewhat difficult given the diversity of curricular and pedagogical requirements posed by the different disciplines, the hodgepodge of standards that currently support medical data formatting and distribution (although see “Proposed Guidelines for EHR Relevance to Computerized Patient Case Scenarios for Education” document for more on the Electronic Health Record (EHR)), and the emergence of simulation-savvy medical professionals who realize profit in authoring and selling case scenarios to colleagues. However, the solution may lie in the inevitable standardization of patient case data and the push for medical curricula that are designed around the use of simulators for training. In addition, as simulation-based training in medicine becomes more common and thus more of a viable commercial enterprise, then a convergence of solutions in this domain becomes inevitable.

•
•
•
•
•

A key R&D goal is to build repositories of VP data that can be used to generate on the fly scenarios that support a variety of teaching and training goals. To reduce the cost of creating these repositories, the goal should be to: a) enable automatic extraction of relevant data from electronic patient records and b) make it easier to create hypothetical VP data designed to address specific teaching or training objectives. This data should be created and stored as potentially reusable and interoperable learning objects, and include metadata that describes the learning object. The patient case data should be separate from the applications through which they are rendered and manipulated, and should be designed such that they are adaptable and interchangeable. Tools and services are required to assist developers in the application of metadata at all levels of content development. These tools need to be customized to meet the needs of various communities of practice.

R&D milestones and subtasks are:

- 3-year milestone:
 - Define work flow and cost issues related to VP data
 - Establish frameworks for consistent nomenclature and vocabularies;
 - Standards and tools that enable “visual cues”, drawn from the clinical examination of the appearance of the patient, as well as vital signs, brief medical history, lab results and ability to reference diagnostic images
- 5- year milestone:
 - Establish standards for nomenclature and vocabularies and define semantics of metadata and implementation guidelines for developers
- 10-year milestone:
 - Tools capable of intelligently generating metadata to support on the fly search/retrieval

(2) VP data search and retrieval services to make it easier to generate scenarios.

Tools are required to build collections of the VP content objects and apply sequencing rules that enable their use in scenarios for delivery as educational and training tools. A special area of required research is the development of tools and/or agents that can do intelligent searching of metadata during authoring of cases and eventually real time for “on the fly” scenario generation.

R&D sub-tasks are:

- 3-year milestone:
 - Salient search attributes for creating content objects defined; prototype search and retrieval tools demonstrated
- 5-year milestone:

- Rules that intelligently guide search and retrieval defined and demonstrated; tools demonstrated that establish connections and connect search engines to repositories of VP data
- 10-year milestone:
 - Robust search methodologies (Google for VP learning objects) available
 - tools and services capable of aggregating learning activities/content based on metadata

(3) Interoperable VP physiological models.

Patient simulators use sophisticated physiological models that are typically tailored to a specific patient case scenario to produce physiological changes in the simulator. Ideally, it should be possible to integrate these physiological models to: 1) produce one-the-fly scenarios, and 2) produce more complex medical cases than are possible today.

Unfortunately, the current state of the commercial marketplace today dictates that different proprietary models, and in some cases proprietary programming languages and patient case building code, prevents interoperability and inhibits a more open approach to creating different Virtual Patient scenarios. Obviously, this can also make life more difficult for the clinician and developer to create their own physiological models, and versions that could run on different simulators.

The reality is that we are decades away from not only understanding all of the physiology and pathophysiology components that operate in a real human, and so even if simplification of model development can reasonably simulate physiologic state, the effects of drug and other interventions, they need to be constrained to address specific learning objectives that account for a majority of human physiology.

This leads to a conundrum. Either the experienced medical instructor has to closely monitor and adjust any particular case to ensure validity and introduce variables that would reflect what would encounter in real world patients, often as a consequence of how the student is manipulating the model, or the physiological objects themselves have to be “intelligent” in the sense that they can be programmed with appropriate parameters to interact and produce output in such a manner that is accurate and informative for instructional purposes. This is often accomplished using object-oriented programming, but the danger is that some unspecified or inaccurately defined variable will guide this kind of ongoing model in a non-physiologic direction.

Thus, hard-coded solutions such as declarative programming provide a safer solution to modeling the various Virtual Patient physiology models. The code is configured as an “If, Then” or “When, If” configuration, with “When modules representing discrete physiological variables such as heart rate. The code is constantly evaluated to see if any “When” module has changed, the process occurring many times per second. Declarative programming can be problematic, not only because it is laborious, but if a key variable is not programmed, such as decreases in respiratory rate following morphine injection, then the resulting patient simulation will be unrealistic. Add to this the explosion of pharmacologic agents that need to be administered, and an

•
•
•
•
•

almost complete lack of knowledge about drug interactions and the summative, antagonistic and/or synergistic interactions they may produce, leaves us a long way from having an ideal library of Virtual Patient models, even if the only goal is only instructional, not predictive.

The ultimate goal for developing methods for sharing physiological models, which preserve the intellectual property rights of the authors/developers, deepened on some standards organization (such as DOD) imposing methods for interoperability. The Federation of American Scientists' Digital Human Project is working to build a collaborative community of information and biological scientists to accelerate development and use of biomedical simulation for research and learning. The project's goal is development of a framework that will permit researchers to collaborate, share their work, and test each other's systems.

Achieving interoperability between simulation systems requires: (1) a common network software architecture with standard protocols that govern the exchange of information about the state of the simulation; (2) a common underlying architecture for maintaining information about the state of the environment related to a particular simulator; and (3) a common representation of the synthetic environment.

Effective combination and reuse of software objects requires precise agreement on the coordinate systems and methods for representing complex geometric objects, the system of units employed, and the exact terminology used to describe objects (or ontology). This last task appears mundane but proves to be extraordinarily difficult since many specialties in the same discipline can use different words to describe the identical object. While work on ontologies is essential, they provide only a structured vocabulary of nouns. Simulations show motion and interaction and require a precise taxonomy of verbs – that is rates of change and flows of charge, chemicals, and bulk materials. They must also show changes in shape and even basic topology of objects.

R&D sub-tasks are:

- 3-year milestone: A defined unified ontology for physiological models; an agreed upon modeling framework defined, with some successful prototypes demonstrated.
- 5-year milestone: Adoption of architectures and standards that support broader development. Established interfaces that enable interoperation of models.
- 10-year milestone: Multi-resolution and hybrid modeling frameworks and toolkits, applied to range of domains.

Table 6 summarizes the R&D topics for automated scenario/case authoring tools.

Table 6. R&D Topics for Automated Scenario/case Authoring Tools

Tasks	Milestones		
	3-years	5-years	10-years
Methods and tools to reduce time and cost to create data for scenarios	<p>Define work flow and cost issues related to VP data</p> <p>Establish frameworks for consistent nomenclature and vocabularies</p> <p>Standards and tools that enable “visual cues”, drawn from the clinical examination of the appearance of the patient, as well as vital signs, brief medical history, lab results and ability to reference diagnostic images</p> <p>Define and prototype a library of mini-context strategies that build on instructional methodology</p> <p>Prototype tools for the construction libraries</p>	<p>Establish standards for nomenclature and vocabularies and define semantics of metadata and implementation guidelines for developers</p> <p>Design/research new methods of developing and using content templates for a wide variety of training and educational needs</p>	<p>Tools capable of intelligently generating metadata to support on the fly search/retrieval</p>
Data standards for automated generation of the patient case presentation	<p>Determine work flow and cost issues related to data</p> <p>Establish frameworks for consistent nomenclature and vocabularies</p>	<p>Establish standards for nomenclature and vocabularies</p> <p>Create tools that reduce author intervention during creation.</p>	<p>Standards and tools enable “visual cues”, drawn from the clinical examination of the appearance of the patient, as well as vital signs, brief medical history, lab results and ability to reference diagnostic images</p>
Content Formats	<p>Research a range of learning content types including simulations, virtual environments, interactive media, collaboration, assessment, etc. Develop taxonomy and examples for each.</p>	<p>Determine authoring requirements and tools for various types that hide technical implementation details</p>	<p>Create tools that seamlessly integrate varied content types for non-technical authors</p>

•
•
•
•
•

Table 6. R&D Topics for Automated Scenario/Case Authoring Tools (continued)

Tasks	Milestones		
	3-years	5-years	10-years
Meta Data for Learning Content	Create tools to relate data from multiple sources that are instructionally sound. Create search tools that work with these data	Create tools that intelligently generate metadata to support on the fly search/retrieval	Develop tools that implement metadata practices for multiple domains
Storage, Search and Retrieval Services	Determine salient search attributes for creating content objects; prototype Develop advanced, intelligent search strategies and engines; build tools that “front end” the engines	Build tools & services that aggregate learning activities/content based on metadata Develop robust search methodologies (Google for learning objects)	Develop rules that intelligently guide search and retrieval; build tools that establish criteria, domain, etc. Develop and refine connections and search engines to repositories of media (scale up)
Frameworks to enable interoperable VP physiological models	A defined ontologies for physiological models 1-3 successful prototypes demonstrated	Adoption of architectures and standards that support broader development Established interfaces that enable interoperable models	Multi-resolution and hybrid modeling frameworks established Toolkits applied to range of domains

5.2 Research Focus Topic #2: Modeling Learners and Intelligent Tutoring

Intelligent tutoring systems (ITS) seek to replicate the experience a learner has with a good human tutor, providing one-on-one instruction. In a landmark series of studies, Bloom and colleagues demonstrated that one-on-one tutoring improved student achievement by two standard deviations over group instruction. (This is often referred to as the “2-Sigma effect”, and is the equivalent of raising the performance of a student from the 50% percentile to the 98th percentile.) Although learning scientists continue to seek to fully understand why such a dramatic difference exist between one-on-one tutoring and group instruction, there is general agreement that: individualization (that instruction can be tailored to the learner’s particular needs), and instructional intensity (the number of interactions between teacher and student during a tutorial are key. There are a number of assessments of ITS that have found that students using cognitive tutors perform about a letter grade (approximately one sigma) better than students who do not (Anderson, Corbett, Koedinger, Pelletier, R. (1995); Koedinger, Anderson, Hadley, Mark, 1997). The approach also appears to produce more motivated students and considerable teacher acceptance (Anderson and Gluck, (2001).

Computer-based, artificial intelligence tutoring technology has been around for several years, especially in military domains where it has matured in a rapid fashion. Research reports support the notion that Intelligent Tutoring Systems (ITS) have tremendous instructional benefits (Ramachandran and Domeshek, 2005):

- Dr. Wes Regian (Air Force Labs) reports average improvement of 1 standard deviation, compared to classroom instruction
- John Anderson of Carnegie Mellon University reports that ITSs:
 - Required 1/3 less instruction time
 - Yields 1 standard deviation performance improvement
- Air Force avionics tutor evaluation
 - 20 hours with tutor = 4 years on-the-job experience
- US Navy reports 1000% increase in tactical experience with Stottler Henke’s Tactical Action Officer ITS

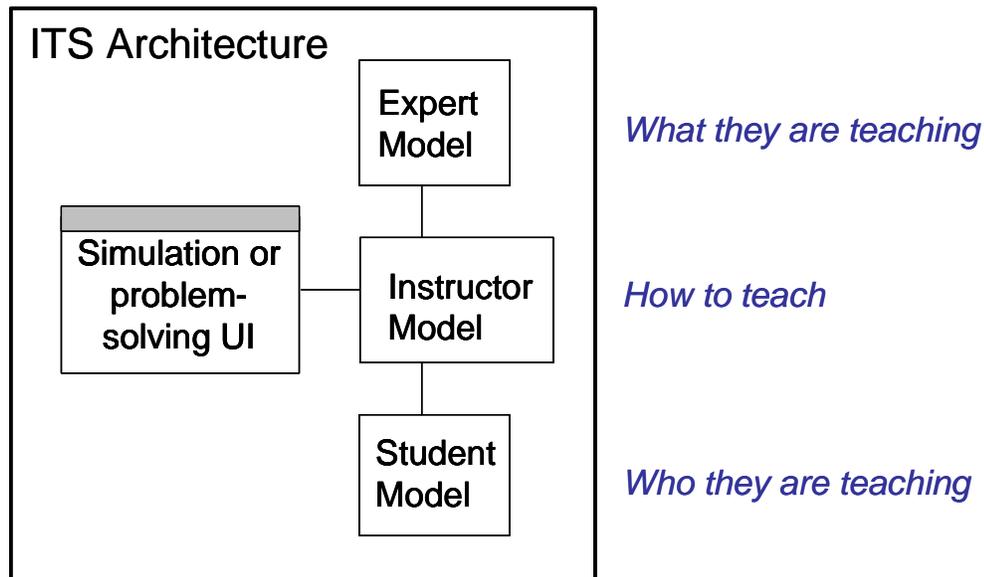
Intelligent tutoring systems (ITS) also have been shown to be effective in tutoring students in various topics including algebra, chemistry, and physics. Anderson and his colleagues have conducted seminal work in intelligent tutoring and have demonstrated its effectiveness in a limited number domains (Anderson, et al., 1999; Corbett, Koedinger and Anderson, 1995).

Pham et al., 2005 compared two intelligent tutors, Rapid Fire / Smart Tutor and the Minimally Invasive Surgery Trainer for Virtual Reality (MIST VR) for laparoscopic

performance improvement. Users of both systems show improvement of laparoscopic skills as measured by paper cutting exercises. Stredney, et al., 2000 integrated an intelligent tutor into a simulation-based trainer for residents in otologic surgery techniques, providing the learners with multiple ways to query and receive information.

ITS are constructed around a cognitive model of the knowledge the learner is acquiring. The cognitive model is used to trace the learner's solution path while solving a complex problem. The ITS recognizes and assesses the learner's competency on each of the cognitive skills. The ITS can provide feedback on each problem solving action and give advice on steps to solving the problem. The assessment is based on the cognitive model that represents individual cognitive skills required to perform a target task. Figure 5 describes the types of information modalities needed by the ITS architecture.

Figure 5. Information modalities needed by an Intelligent Tutoring System architecture. Source: Ramachandran; Eric Domeshek, 2005.



Ramachandran and Domeshek (2005) describe the key elements of an intelligent tutoring system as shown in Table 7.

Table 7. Elements of Intelligent Tutoring Systems. Source: Ramachandran and Domeshek (2005).

Important Elements of Intelligent Tutoring Systems
<p>1. Performance Assessment and Scaffolding</p> <ul style="list-style-type: none">• Evaluation of student performance under free-play conditions generally challenging<ul style="list-style-type: none">• Rules, Constraints• Finite State Machines• Intelligent Pattern Matching• Natural Language Processing, Latent Semantic Analysis• A student model is a system's evolving longer-term picture of what a student has experienced, knows, and can do, possibly combined with a component that assesses learning preferences and style.• Informs decisions about individualized scenario choice and sequencing• May also inform decisions about individualized instructional interventions and coaching
<p>2. Coaching and After-action Debriefing</p> <ul style="list-style-type: none">• Approaches include:<ul style="list-style-type: none">• Hinting (proactively or reactively)• Immediate correction (possibly with instruction)• Deferred commentary (at natural break, or when consequences manifest)• Reformulate problem or environment to simply/focus task• Appropriate techniques depend on nature of the problem domain, simulation, proven instructional practices, and available technology• Opportunity to review with students their performance in a simulation<ul style="list-style-type: none">• Dynamic report card• Linkage to explanations and reference materials• Relationship to prior performance• Scenario playback• What-if scenarios• Socratic dialogs
<p>3. Intelligent Scenario Control</p> <ul style="list-style-type: none">• Many possible attributes ITS could seek to control:<ul style="list-style-type: none">• Which scenario to present when• How major events of scenario unfold• How entities within the scenario behave• How tutor responds in context of scenario (coaching)• How tutor responds at close of scenario (after action)• Cognitively inspired models of behaviors for virtual entities (e.g. hybrid reactive-deliberative architectures) have proven useful in many applications

•
•
•
•
•

Building a good ITS is difficult, as it requires knowledge of the subject matter, a good understanding of the prior abilities of the students who will use the system, and generating and maintaining a continuous/dynamic model of profile of the learner to can be used for analysis, feedback, coaching during instruction. Once collected, learner performance data must be interpreted in a manner that allows conclusions about learner mastery to be drawn. This implies two capabilities: first is the ability to interpret dynamically collected performance data in a manner that allows meaningful diagnosis to occur, and second is to develop a means to compare this “observed” performance to an expert standard. The ability of the ITS to communicate the contents of the expert model to the learner is crucial. In this regard, issues such as the quality of help and error messages are of interest. The intelligent tutor must accurately diagnose what the student does and does not know and deliver feedback and/or remediation appropriately.

The learning environment must incorporate appropriate support, or scaffolding, for learners as a means to guide them through the learning process, including adjusting task difficulty to the learner’s current level of ability, restructuring the task to supplant knowledge, and providing alternative assessments and worked examples. The notion that a simulated learning environment must incorporate appropriate scaffolds for learners as a means to guide them through the learning process has received some attention in the literature. For example, Bransford, Brown and Cocking (1999) discuss the use of technology to scaffold experience. They use the analogy of “training wheels” as a means to explain how computerized tools can be used to support learning that students would otherwise be unable to accomplish. Hannafin, Land and Oliver (2000) describe several methods and mechanisms for scaffolding, including conceptual, metacognitive, procedural, and strategic. Jonassen (2000) also discusses several types of scaffolding for learning: 1) adjusting task difficulty to the learner’s current level of ability, 2) restructuring the task to supplant knowledge, and 3) providing alternative assessments and worked examples. In addition, Reiser et al. (2001) have built software tools to scaffold student inquiry in science. All of these techniques can be supported by technology in the learning environment; the nature and applicability of them requires further inquiry.

In addition, coaching has been implicated as means to enhance learning for understanding – to give adaptive feedback based on task analysis results indicating the manner in which expertise develops in this domain. All of these techniques can be supported by technology in the learning environment; how to most effectively use these approaches requires further study.

According to Anderson and Schunn (2000), there are three key issues related to design of intelligent tutoring systems. First is the accuracy of the underlying cognitive or expert model (which relates back to the issue of cognitive task analysis discussed in Section 5.1.1. Development of the expert model requires specification of the constructs (skills, knowledge, abilities) to be measured. This requires decomposition (analysis) of the content/job/performance domain into its constituent knowledge and skill components. Such decomposition results in lists, clusters, and hierarchies of skill components (or learning/assessment objectives) at various levels of granularity. The more fine-grained the decomposition, the more test items, or responses within larger assessment tasks, can be targeted at particular subskills, and the more specific

the diagnosis of knowledge gaps. The more the decomposition of expertise is based on a cognitive tasks analysis, the greater the validity, efficiency, and effectiveness of the resulting assessment and learning (Clark and Estes, 1996; Clark, 2003; Schraagen, Chipman, and Shalin, 2000).

The problem is that there are many different taxonomies and models of cognitive processes, structures, knowledge types, and learning objectives that can be used for classification of component knowledge and skills. The existence and use of so many models for classifying cognitive processes, types of knowledge and learning outcomes makes it difficult to agree on general components of performance; agree on reusable components, metadata, and templates for assessment development; and standardize the interpretation of assessment data from multiple sources (FAS, 2003). While some work has been done in this area, much work remains. Cao, et al., 1999 have developed an analytical framework based on a hierarchical decomposition for studying motor task performance during laparoscopic procedures among surgeons. Cristancho, et al., 2006 extended this model to model the cognitive aspects and describe the event sequences. Heinrichs et al., 2004, developed a structured vocabulary defining fundamental target skills familiar to surgeons to facilitate communication among surgeons and the engineers developing surgical simulators.

The long-term goal is to develop common frameworks for competency models. This could in turn drive the development of tools and software for cognitive task analysis, including generation of content/skill maps and learning objectives in a form that would enable the automated generation of tasks to elicit and measure those skills. Automated tools would greatly reduce the cost of developing cognitive tutors.

According to Anderson and Schunn (2000), the second key issue for intelligent tutoring systems is the ability of the tutor to communicate the contents of the expert model to the student is crucial. In this regard, issues such as the quality of help and error messages are important. Third, the tutor must accurately diagnose the learner's mastery—what the student does and does not know—and deliver feedback and/or remediation appropriately.

The key research tasks for modeling learners and intelligent coaching are:

- Common frameworks to define competency models
- Automated tools collecting performance data and monitoring performance
- Automated tools for translating cognitive task analysis data into diagnostic models
- Guidelines and tools to optimize the introduction, format, timing and fading of assistance to the learner
- Coaching strategies that dynamically adjust according to learner achievement
- Automated processes for generating and presenting on-line feedback that is sensitive to the task and to learners

Table 8 provides the key research tasks.

•
•
•
•
•

Table 8. Research Topics for Modeling Learners and Intelligent Tutoring

Tasks	Milestones		
	3-years	5-years	10-years
Performance Modeling	Mapping of representative domain-specific and domain-general content/ competency models	Agreement from major constituent groups on core model, terms and relations Illustration in multiple domains	Widespread use of common framework and language to discuss and define components of expertise and competence Reporting of all assessment data in all domains in terms of common constructs and framework
Learner Monitoring and Instrumentation	Identification of data to be collected for assessing learner performance	Guidelines for incorporating keystroke data, eye track data, speech, gesturing and facial expressions (if applicable) into a performance monitoring system	Guidelines for instrumenting learning systems to collect performance data dynamically Automated tools for monitoring learner performance
Diagnostic Models	Empirical data to support approaches to diagnostic modeling that are specific to the task/domain	Demonstration of various modeling formalisms linked to tasks/domains to diagnose performance	Automated tools for translating cognitive task analysis data into diagnostic models Methods to dynamically apply diagnostic models to on-going performance
Scaffolding	Empirical results linking learner characteristics to the need for scaffolding Empirical studies regarding successful methods to scaffold learning	Demonstration of techniques that optimize the introduction, format, timing and fading of scaffolding in the learning environment Guidelines for scaffolding according to task type and learner characteristics	Automated tool that adjusts scaffolding strategies automatically as a function of learner characteristics and on-going performance
Coaching	Empirical results indicating the effectiveness of various coaching strategies	Empirical results linking coaching strategies to learner characteristics and task features	Guidelines for coaching strategies adjusted for learner characteristics and task type Demonstration of automated coaching strategies that dynamically adjust according to learner achievement
Feedback and Remediation	Empirical data regarding to timing and specificity of on-line feedback Empirical data regarding the structure and content of post-exercise feedback	Guidelines for implementing feedback strategies across tasks and learners	Automated process for generating and presenting on-line feedback that is sensitive to the task and to learners

5.3 Research Focus Topic #3: Debriefing and Assessment

A tremendous amount of work has been done regarding how to provide feedback to enhance learning and transfer. Feedback is the central mechanism by which learners can regulate their own performance and understand how to improve. According to Bransford, Brown and Cocking (1999), opportunities for feedback should be frequent and/or continuous. It has also been argued that feedback should help learners to understand how to change their performance in order to improve (i.e., simply providing learners with knowledge of results is insufficient) (Bransford, Brown and Cocking, 1999). Recently, it has been argued that feedback strategies that promote deeper processing are superior to achieve transfer of learning. For example, intermittent feedback may help trainees to be less dependent on continuous reinforcement (Schmidt and Bjork, 1992).

Assessment is defined as the measurement of learners' knowledge and skills, as well as measurement of other personal characteristics that influence learning and performance (Snow and Lohman, 1989; Kyllonen, 2000). Assessment is a process of reasoning from evidence (Mislevy, 1994, 1996) to determine a learner's competence and is dependent on the types of evidence or observations and the types of tools that are available for interpreting the evidence (or data). The constructs to be assessed can go beyond the individual learner to include aspects of the learning and transfer environment. Assessment can occur before, during or after a learning opportunity and can include retention and transfer.

Evaluation and assessment provide the data for decision-making regarding:

- What knowledge and skill gaps of individuals and groups need to be targeted with instruction;
- What feedback, guidance, and learning resources to provide individual learners during the learning process;
- What progress the learner is making in each domain area or scale;
- Which educational programs or components of educational programs are ineffective or inefficient and need modification;
- Who is competent to perform particular tasks;
- Who to select or promote into particular jobs.

These decisions are only as valid as the data and data interpretations that are available. Ignore a critical variable or use an inappropriate data source and a misguided surgeon could make a crucial misdiagnosis, Ignore performance data altogether and you'll never know if your education and training programs are effective, or if your physicians and nurses could reduce errors, work more effectively, improve processes faster, and innovate more frequently. Forget to measure and address variables related to motivation and you may end up with low productivity or high turnover.

5.4 Feedback

Feedback and guidance are essential components of a learning environment so that errors in performance can be pointed out and corrected and the learner can proceed to

•
•
•
•
•
•
•
•

mastery. There are many dimensions of feedback and guidance that can be varied, for example, timing, content, amount, specificity, medium, and control. Researchers have already studied these feedback variables in the context of computer-based instructions and many reviews of the literature are available (e.g., Kluger and DeNisi, 1996; Salas and Dempsey, 1993).

Rules are also needed for selecting what the target of feedback will be; for example, when a learner makes an error, what set of values for the content, motivation, and metacognitive components of the learner model would trigger feedback that targets motivation as the first option to try. Once we have established rules for feedback decisions, we need software that allows an author to specify rules for triggering particular types of feedback. Authoring software also needs to allow for the entry of feedback segments that can intelligently and dynamically be pieced together, or presented in a variety of media, for example text, or spoken by a character.

Personalization of content can occur at a macro or micro level; a multilevel learner model will permit either or both depending on the granularity of the content. One area requiring research is the level of granularity of content required to take advantage of more granular learner models.

Validating rules for feedback and personalization of content will encompass many studies where rules for triggering feedback based on states of the learner mode are varied, aspects of feedback itself are varied, and the resulting impact on learner performance is compared. Rules and strategies that reduce time to learn and increase mastery will be deemed the best. Feedback and guidance in the form of hints and on demand access to help should also be investigated. This task will generate rules for generating multi-faceted feedback and personalization of content (e.g., Hsieh and O'Neil, 2002; O'Neil, Chuang, and Chung, in press).

Developing interfaces for authoring feedback mechanisms will focus on developing authoring tools components and interfaces for specifying complex rules and content for feedback. These tools should incorporate the rules that were generated by the research in the first task (e.g., Muraida, Spector, O'Neil, and Marlino, 1993; O'Neil, 1979a, 1979b; O'Neil, and Baker, 1993, 1997; O'Neil, Mayer, Herl, Niemi, Olin, and Thurman, 2000; O'Neil, Wang, Chung, and Herl, 2000).

5.5 Key Elements of Assessment

Four decades of research in the cognitive sciences has advanced the knowledge base about how people develop understanding, how they reason and build structures of knowledge, which thinking processes are associated with competent performance, and how knowledge is shaped by social context. Assessment is a tool for observing learners' behavior and producing data that can be used to draw reasonable inferences about what learners know. The process of reasoning from evidence has been described as the *assessment triangle*, comprised of three elements: a theory of how learners represent knowledge and develop competence in the subject domain (cognition); a set of beliefs about the types of observations that provide evidence of the learner's competencies; and an interpretation process for making sense of the

evidence (NRC, 2001). Each of the three elements of the assessment triangle must connect to each of the other elements in a meaningful way.

The constraints of current traditional testing methods make it difficult to assess many aspects of cognition and expertise. Fortunately, emerging technologies are making it possible to assess a wider range of cognitive competencies by presenting complex, open-ended problems and simultaneously collecting information on how learners go about solving them. At the same time there have been significant developments and measurement methods and theory. A wide array of statistical measurement methods is currently available to support the kinds of differences that cognitive research suggests.

Technology offers opportunities to strengthen the cognition-observation linkage of the assessment triangle by enabling the design of situations that assess a broader range of cognitive processes than was previously possible, including knowledge-organization and problem-solving processes that are difficult to assess using traditional, paper-and-pencil assessment methods. Technology makes it possible to analyze the sequences of actions learners take as they work through a problem and to compare these sequences of actions against models of knowledge and performance associated with different levels of expertise. Technology can make possible stronger observation-interpretation linkages through improved analysis and scoring methods.

Developing real-time flexible reporting systems focuses exclusively on reporting of learner model data for different audiences. Reporting mechanisms, layouts, and query interfaces of different types can be created and compared for usability by the audiences.

Research is needed to identify and test the relative utility, efficiency, and validity of both obtrusive and unobtrusive indicators of metacognitive and motivational variables in simulation-based learning environments. We ultimately want validated guidelines for the most appropriate learner model and measurement techniques depending on budget, context, and purpose. An approach to cost-benefit studies of learner models with various types of measures and levels of granularity might be to systematically eliminate components or layers of detail in a learner model, and examine the effect on other variables in the model and on the instructional modifications and actual learner performance.

Evaluation data can also be used to assess how well systems, rather than individual or teams, are performing. Schaefer, Dongelli, and Gonzalez (1998), used automated data collection from mock codes using patient simulators to identify hospital system deficiencies. Their study involved collecting data from multiple mock codes over a 6 month period in a variety of settings. The resulting data was used to identify and correct system errors to improve the overall health system. Youngblood, et al 2005 studied the transfer of surgical trainees' skills acquired on surgical simulators to the operating room setting to compare the effectiveness of two laparoscopic surgery simulators in order to aid in selection of appropriate training methodologies.

Dillon et al., 2004 in a discussion of the future of simulation for medical licensing that, note that simulation has the potential to enable more precise measurement of areas already addressed, but perhaps more importantly the potential to evaluate areas

•
•
•
•
•
•
•
•
•

not currently assessed because of practical or psychometric difficulties and to assess the performance of healthcare teams. They state:

There are also a number of new technologies that could eventually find their way into high stakes medical licensure examinations. Virtual reality and haptic feedback trainers have been used to train physicians in areas such as minimally invasive surgical procedures and vascular interventions (Nackman, Bermann and Hammond, 2004). Unfortunately, the cost for these systems can be high, especially if they are to be used for large scale assessments where many thousands of examinees must be tested. More importantly, these technologies generally target very specialized skills. From a certification or licensure perspective, where the focus centers on measuring fundamental skills as reliably as possible, it may be inefficient and impractical to use these systems. Additional research will be necessary before these types of trainers can be incorporated in high stakes summative assessments.

Life size mannequins (integrated simulators) with realistic airway and cardiovascular attributes have been used to train physicians and other healthcare professionals. For medical licensing examinations, these simulators suffer from some of the same drawbacks as virtual reality and haptic feedback trainers. They are costly and generally target more specialized skills. Moreover, while a number of scoring systems have been developed (Boulet , Murray, Kras, et al.,2003; Morgan, Cleave-Hogg, DeSousa S, et al., 2004) they have yet to undergo the scientific scrutiny that has taken place for SP assessments (Norcinim Boulet, 2003). Nevertheless, as the cost of these mannequins declines, and additional psychometric studies are completed, they could have a unique role within the licensure process, especially for higher order skills.

Table 9 presents the R&D tasks and associated milestones:

- Models and guidelines for personalizing feedback based on the model of the learner
- Interfaces for authoring feedback mechanisms
- Standard data structures and transfer protocols to support combining and reporting assessment data, for the user and the learning systems

Table 9. Research Topics for Debriefing and Evaluation

Tasks	Milestones		
	3-years	5-years	10-years
Validate rules for feedback and personalization of content	<p>Synthesis of existing research on feedback</p> <p>Results of studies on benefits of adding feedback based on motivational and metacognitive states</p> <p>Results of studies that compare mechanisms for <i>triggering</i> feedback/ hints/guidance/content adaptation</p>	<p>Results of empirical studies that compare methods for <i>delivering</i> feedback/scaffolding/ coaching (intelligent agents, multimedia etc.)</p> <p>Results of studies of alternative <i>timing</i> of feedback</p> <p>Results of studies that compare levels of <i>specificity</i> of feedback</p> <p>Decision-aids/rules for personalizing content</p>	<p>Demonstration of increased effectiveness of instruction that incorporates standard rules for feedback and personalization</p> <p>Demonstration of decrease in time to learn when content is personalized</p> <p>Demonstration of additional benefits of motivational/metacognitive modeling over domain knowledge modeling</p>
Develop interfaces for authoring feedback mechanisms	<p>Identification of best authoring tools for rule-based feedback and personalization</p> <p>Demonstration projects</p>	<p>Rule-based feedback authoring incorporated into authoring tools</p>	<p>Validation of automated management of feedback in simulations</p>
Develop software applications for data management and sharing	<p>Identification of systems that need to share data</p> <p>Data structures and application program interfaces (APIs) for transfer of data</p>	<p>Demonstration projects</p>	<p>Widespread use of standard data structures and transfer protocols</p>
Cost/benefit studies	<p>Results of studies that compare costs of development, delivery and validity/utility of data from different measurement methods</p>	<p>Results of studies that examine the cost/benefit of additional levels of granularity in learner models</p>	<p>Decision-aids for choosing different types of measurements and level of granularity based on context, budget and purpose</p>

6 Summary and Implications of the Virtual Patient Research Roadmap

All of medical and allied health education is undergoing profound transformation, including EMS, nursing and the medical school. Decreased income and reimbursement, increased liability coupled to medical error, personnel shortage within domains such as nursing, increased patient assertiveness about their treatment, as well as decreased federal spending on healthcare, are all forcing the system of healthcare delivery in the United States to change. In addition there has been a growth in the number of simulation centers.

6.1 Curriculum Redesign to Integrate New Learning Technologies

One example of change is the proposed new curriculum for the University Of Pittsburgh College Of Medicine. Their approach was to build around new approaches to learning enabled by using new technologies in simulation and information management. Among the goals were to:

- Show accurate 3D simulations of human anatomy accurate that represent not only gross structure but also show systems in operation (blood and lymph flows, muscle and tendons moving, digestive systems in operation) and be adaptable to show the wide range of variations in normal anatomy as well as pathologies.
- Provide vividly and accurate simulations of the operation of cells again showing not only static forms but active signaling and interaction
- Allow easy changes of scale so that the function of systems and pathologies can be explored at the levels of molecules, cells, or gross anatomy.
- Allow simulated operation of sophisticated imaging and laboratory equipment
- Allow interaction with simulated patients, including plausible conversations about symptoms
- Provide repeated practice in general and organ system-focused physical examinations
- Allow individuals and groups to practice tasks requiring collaboration and communication.
- Provide powerful question management tools linking the learner to the literature and to human experts with knowledge of the subject, the context of the question, and the background of the individual asking the question
- Provide practical experience in using new information tools for “just in time” learning that will be needed in practice

The Curriculum Committee of UPMC listed a number of design goals for its work including:

- The curriculum should encourage methods of instruction that foster active learning and should create an intellectually stimulating environment.

- The curriculum should encourage methods of instruction that foster active learning and should create an intellectually stimulating environment.
- The curriculum should emphasize, reward and facilitate the teaching of medical students by providing resources, including both educational tools and educational expertise.
- Clinical exposure should be introduced actively and as early as possible
- Use problem based learning strategies

The Brigham and Women's Hospital and Massachusetts General Hospital Harvard Affiliated Emergency Medicine Residency program has begun a full-scale curriculum re-design fully integrating medical simulation. Each module of the emergency medicine core curriculum was reviewed and learning approaches selected based on specific learning objectives using one of four approaches: case-based seminars with pre-assigned reading; instructor led problem-based learning using micro-simulators; partial task simulators and high-fidelity human patient simulators (Pozner, et al., 2005).

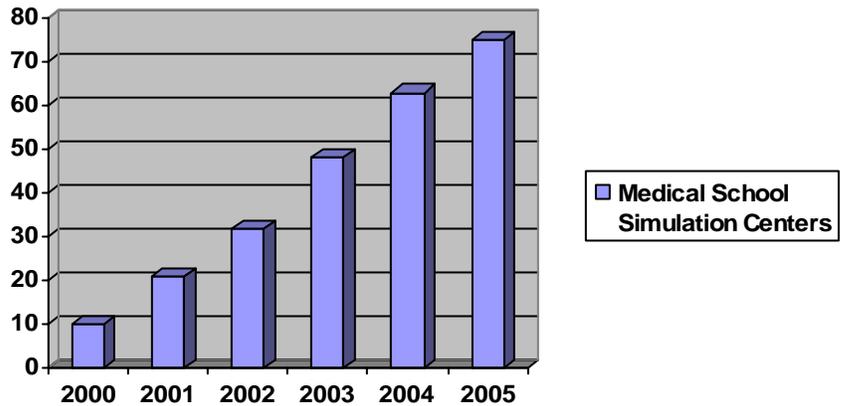
The University of Maryland School of Nursing in the Clinical Simulation Labs (CSL) and the Clinical Education and Evaluation Lab (CEEL), a joint venture between the School of Nursing and Medicine uses a blended learning approach that combines clinical simulation and standardized patients to provide the opportunity for students to develop decision-making skills prior to entering the clinical setting (Spunt and Schaivone 2005).

6.2 Growth in Simulation Centers

There has been tremendous growth in the number of simulation centers, both in medical schools (primarily anesthesiology, although also training in surgery, pharmacology, and nursing, and for training and assessment of undergraduates and graduates), and in the military for medic and EMS training, since 2000. For example, the WISER (Peter M. Winter Institute for Simulation, Education and Research) at the University of Pittsburgh Medical College, during 2004-2005, trained over 8,000 individuals in 10,000 full scale simulations, offered 55 courses in patient simulation, and engaged 25 Course Directors and 150 facilitators. Figure 6 shows the growth of U.S. based medical school simulation centers over the past 5 years. Approximately three quarters of the medical schools in the United and Canada have dedicated medical simulation laboratories.



Figure 6: Growth of Simulation Centers in U.S. Medical Schools.
 Source: U.S. Department of Health and Human Services



In addition to the growth in the number of simulation centers, the community of researchers and practitioners has been growing as is evidenced by an increase in conferences and other convening activities, and highlighted in the following paragraphs.

Advances in Medical Simulation (AIMS) is a coalition of individuals and organizations committed to increasing awareness of medical simulation and providing a community-wide message regarding the benefits that medical simulation provides. The organization’s goal is to engage and further develop the medical simulation community and to secure resources to further medical simulation training, research and the deployment of simulators and simulation tools.

The Society for Simulation in Healthcare (SSH), was established in January 2004 to represent the rapidly growing group of educators and researchers who utilize a variety of simulation techniques for education, testing, and research in health care, emphasizing anesthesiology training. SMS is a broad-based, multi-disciplinary, multi-specialty, international society with ties to all medical specialties, nursing, allied health paramedical personnel, and industry. A major venue for advancing simulation in medicine is the annual International Meeting for Simulation in Healthcare (formerly IMMS).

Medicine Meets Virtual Reality (MMVR) is an annual meeting that showcases innovative research on information-based tools for clinical care and medical education. Research presented at the meeting includes a range of medical simulation work and enabling technologies, including Haptics, tissue modeling, imaging tools, data visualization and fusion, robotics, and tools for medical decision-making. The meeting participants

generally include broad segment of the medical community, including:

- Physicians, surgeons, and other healthcare professionals interested in emerging and future tools for diagnosis and therapy
- Educators responsible for training the next generation of doctors and scientists
- Computer technologists designing systems for gathering, processing, and networking medical intelligence
- IT and medical device engineers who develop and market state of the art imaging, simulation, robotics, and communication tools
- Military medicine specialists addressing the challenges of warfare and defense health needs
- Biomedical futurists and investors who need to understand where medicine is headed

6.3 The Challenge: Integration of Learning Sciences into the Medical Simulation Curriculum

For the Virtual Patient Research Roadmap, we concentrated on three topics that were felt to be most germane to simulation-based education, based on the student's progress:

- Automated patient case generation
- Learning modeling and Intelligent Tutoring
- Feedback and assessment

In an earlier section of the report, we have detailed 1, 3 and 5 year research agendas for each of these topics. Our operating hypothesis has been that implementation of simulation training in medicine will benefit from lessons learned in the learning sciences and other domains such as flight and defense simulation.

However, not all schools of medicine and allied health will be as open to the integration of new simulation and information technologies into their curricula. Thus, schools such as the University Of Pittsburgh Of Medicine will serve as both test beds and if successful, exemplars of how to do it the right way. But we still have a long way to go to convince the medical establishment of the utility of these new learning approaches and their accompanying technological foundations.

For example, at a recent meeting of the 2006 Society for Medical Simulation (SMS mostly attended by anesthesiologists), the future vision seemed to still rely heavily on the usual constrained methods of education and primarily subject methods of assessment. The bulk of emphasis from both the SMS and from the Association of American Colleges (AAMC) vision of the future of learning and assessment in

•
•
•
•
•
•
•
•
•
•

medicine is still based on old-fashioned methods. For example, from the recent publication, AAMC Project on the Clinical Education of Medical Students:

“Moreover, the traditional emphasis upon standardized written assessment in determining medical student professional development fosters knowledge rather than a skill-based paradigm in undergraduate medical education. Not only does a multiple choice test, for example, seem more “objective”, it is also a more efficient and less expensive way to evaluate student achievement than observing and assessing individual skill learning outcomes.”

The inherent weaknesses of the multiple-choice exam have been well documented elsewhere (REF). Other include recommendations for evaluation from peers and senior medical personnel using tools such as 360 degree analysis and personal recommendations for obtaining privileges.

Yet in the final analysis, maybe we can agree with the AAMC, who quoted Sir William Osler in their report:

“In what may be called the natural method of teaching, the student begins with the patient, continues with the patient, and ends his studies with the patient, using books and lectures as tools, as means to an end. The student starts, in fact, as a practitioner, as an observer of disordered machines, with the structure and orderly functions of which he is perfectly familiar. Teach him how to observe, give him plenty of facts to observe and the lessons will come out of the facts themselves. For the junior student in medicine and surgery it is a safe rule to have no teaching without a patient for a text, and the best teaching is that taught by the patient himself. The whole art of medicine is in observation, as the old motto goes, but to educate the eye to see, the ear to hear and the finger to feel takes time, and to make a beginning, to start a man on the right path, is all that we can do. We expect too much of the student and we try to teach him too much. Give him good methods and a proper point of view, and all other thing.”

The only way in which students can be exposed to the greatest number of different patients, anomalies, and varieties of pathologies is through the use of learning through the use of simulated patients. For example, in medical school, interns can never see all of the possible cases that might encounter during their career, yet a properly configured and validated simulator could expose to this range of variability at their leisure. Now that the technology is in place, work must begin to work on the vision in this Virtual Patient Research Roadmap.

7 REFERENCES

- Advanced Distributed Learning. www.adlnet.org [Last accessed December 5, 2005.]
- Andrews, D.H., Carroll, L.A., Bell, H. (1995). The Future of Selective Fidelity in Training Devices. Armstrong Laboratory Technical Report AL/HR-TR-1995-0195.
- Anderson, J. R. and Gluck, K. (2001). What role do cognitive architectures play in intelligent tutoring systems? In D. Klahr and S. M. Carver (Eds.) *Cognition and Instruction: Twenty-five years of progress*, 227-262. Erlbaum.
- Anderson, J. R., Corbett, A. T., Koedinger, K., and Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of Learning Sciences*, 4, 167-207.
- Anderson, J. R. and Schunn, C. D. (2000). Implications of the ACT-R learning theory: No magic bullets. In R. Glaser (Ed), *Advances in instructional psychology: Educational design and cognitive science*, Vol. 5. (pp. 1-33). Mahwah, NJ: Lawrence Erlbaum Associates.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191-215.
- Benton, C.J., P. Corriveau, J.M. Koonce and W.C. Yire (1992) Development of the basic flight instruction tutoring system (BFITS). Air Force Systems Command. AL-TP-1991-.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- 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.
- Boldovici, J.A. (1987) Simulator motion. Army Research Institute. TR961 ADA257683.
- Boulet JR, Murray D, Kras J, et al. Reliability and validity of a simulation-based acute care skills assessment for medical students and residents. *Anesthesiology* 2003;99:1270–80.
- Bransford, J. D., Brown, A. L., and Cocking, R. R. (Eds.) (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.
- Cao, C., Mackenzie, C., Ibbotson, J., Turner, L., and Blair, N., (1999). Hierarchical decomposition of laparoscopic procedures. *MMVR*, IOSPress, pp. 83-89.
- Caro, P. and R. Isley (1966) Helicopter trainee performance following synthetic flight training. *Proc. Southwestern Psychological Association*.

•
•
•
•
•
•
•
•

Caro, P., R. Isley and O. Jolley (1973) Research on synthetic training: Device evaluation and training program development. HumRRO Division No. 6. U.S. Army Research Institute for the Behavioral and Social Sciences.

Caro, P., R. Isley and O. Jolley (1975) Mission suitability testing of an aircraft simulator. HumRRO-TR-75-12. U.S. Army Research Institute for the Behavioral and Social Sciences.

Caro, P., W. Corley, W. Spears and A. Blaiwes (1984) Training effectiveness evaluation and utilization demonstration of a low-cost cockpit procedures trainer. Naval Training and Equipment Center, NAVTRAEQUIPCEN78-C-0113-3.

Carretta, T.R. and R.D. Dunlap (1998) Transfer of Effectiveness in Flight Simulation: 1986 to 1997. U.S. Air Force Research Laboratory, NTIS ADA362818.

Connolly, T.J., B.B. Blackwell and L.F. Lester (1989) A simulator-based approach to training in aeronautical decision making. *Aviation, Space and Environmental Medicine* 60:50-52.

Clark, R. C. (2002). The new ISD: Applying cognitive strategies to instructional design. *Performance Improvement*, 41, 8-15.

Clark, R., and Mayer, R.E., (2003), *e-Learning and the Science of Instruction*. San Francisco, CA. Pfeiffer.

Clark, R. and Wittrock, M. (2000). Psychological principles in training. In S. Tobias and J. D. Fletcher (Eds.), *Training and retraining: A handbook for business, industry, government, and the military* (pp. 51-84). New York, NY: Macmillan..

Cognition and Technology Group at Vanderbilt (CTGV) (2000). Adventures in anchored instruction: Lessons from beyond the ivory tower. In R. Glaser (Ed), *Advances in instructional psychology: Educational design and cognitive science*, Vol. 5. (pp. 35-99). Mahwah, NJ: Lawrence Erlbaum Associates.

Cognitive Task Analysis (2000) Jan Maarten Schraagen (Editor), Susan F. Chipman (Editor), Valerie L. Shalin (Editor). Lawrence Erlbaum Publishing, MahWah, NJ.

Cyrus, M.L. (1978) Motion systems role in flight simulation for flying training. Air Force Human Resources Laboratory. AFHRL-TR-78-39, ADA059744.

Digital Human White Paper. Howell, K., Kelly, H., Higgins, G., Federation of American Scientists. March 2006. www.FAS.org/digitalhuman [Last accessed March 28, 2006.]

Dillon GF, Boulet JR, Hawkins RE, et al. Simulations in the United States Medical Licensing Examination (USMLE). *Qual Saf Health Care* 2004; 13(suppl 1):i41-i45.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155-170.

Dixon, K.W. and D.G. Curry (1990) Weapons delivery training: Effects of scene content and field of view. Operations Training Division, Air Force Human Resources Laboratory. AFHRL-TP-88-29, ADA227968.

Dixon, K.W., E.L. Martin, V.A. Rojas and D.C. Hubbard (1990) Field-of-view assessment of low-level flight and airdrop in the C-130 weapon system trainer (WST). Operations Training Division, Air Force Human Resources Laboratory. AFHRL-TP-89-9, ADA218504.

Dohma, J.A. (1995) The military quest for flight training effectiveness. In Larsen, W.E., R.J. Randle, Jr. and L.N. Popish, Vertical Flight Training, pp. 87-130, NASA and DOT.

Durall, E., W. Spears and W. Prophet (1978) Reduced physical fidelity training device concepts for army maintenance training. Naval Training Equipment Center. PM Trade-SE-78-0002.

Edwards, B. (1986) Low-cost avionics simulation for aircrew training. Air Force Human Resources Laboratory. AFHRL-TR-85-38.

Federation of American Scientists (2003). The Learning Federation Project Learning Science and Technology R&D Roadmaps: Integration Tools for Building and Maintaining Advanced Learning Tools (2003).

<http://www.fas.org/main/content.jsp?formAction=297&contentId=69>. [Last accessed March 28, 2006.]

Gaba DM, and A. DeAnda (1988) A comprehensive anesthesia simulation environment: re-creating the operating room for research and training. *Anesthesiology*. Sep;69(3):387-94.

Gallagher, A., Ritter, E., Champion, H., Higgins, G., Fried. M., Moses, G., Smith, C., and Satava, R. (2005) Virtual Reality Simulation for the Operating Room; Proficiency-Based Training as a Paradigm Shift in Surgical Skills Training. *Annals of Surgery*. 241(2):364-372. February.

Graesser, A., VanLehn, K., Rose, C., Jordan, P., Harter, D. (2001) Intelligent tutoring systems with conversational dialogue. *AI Magazine*, 2001.

Guinn, F., Hubal, C., Pope, R., Stanford, M., and Lamm-Weisel, D. (2002), "JUST-TALK: An Application of Responsive Virtual Human Technology" in Proceedings of the Interservice/Industry Training, Simulation and Education Conference, Orlando, FL, USA.

Guckenberger, D., K.C. Uliano and N.E. Lane (1993) Teaching high performance skills using above real-time training. NASA Contractor Report No. 4528.

Hannafin, M. J., Land, S., and Oliver, K. (1999). Open learning environments: Foundation, methods, and models. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: A new paradigm of instructional theory*, Vol. II. (pp. 115-140). Mahwah, NJ: Lawrence Erlbaum Associates.

•
•
•
•
•
•
•
•

Hays, R.T. and M.J. Singer (1989) *Simulation fidelity in training system design*. New York. Springer Verlag.

Hays, R.T., J.W. Jacobs, C. Prince and E. Salas (1992a) Flight simulator effectiveness: A meta-analysis. *Military Psychology* 4(2):63-74.

Hays, R.T., Jacobs, J.W., Prince, C. and Salas, E. (1992b). Requirements for Future Research in Flight Simulation Training: Guidance Based on a Meta-Analytic Review. *The International Journal of Aviation Psychology* 2 (2):143-158.

Hettinger, L.J., W.T. Nelson and M.W. Haas (1994). Applying virtual environment technology to the design of fighter aircraft cockpits: Pilot performance and situation awareness in a simulated air combat task. In: *Proceedings of the 1994 Human Factors and Ergonomics Society Meeting*, (pp. 115-118), Santa Monica, CA: Human Factors and Ergonomics Society.

Hettinger, L.J., W.T. Nelson, B. Brickman, C. Roumes and M.W. Haas (1995). Assessing human performance in virtual environments. In *Proceedings of the Eighth International Symposium on Aviation Psychology* (pp. 170-176). Columbus, OH.

Higgins, G.A., D.A. Meglan, R. Raju, J.R. Merrill and G.L. Merrill (1997a) Teleos: Development of a software toolkit for authoring virtual medical environments. *Presence* 6(2), 241-252.

Higgins, G.A., G.L. Merrill, L.J. Hettinger, C.R. Kaufmann, H.R. Champion and R.M. Satava (1997b) New simulation technologies for surgical training and certification: Current status and future projections. *Presence* 6(2), 160-172.

HEAL: <http://www.healcentral.org/index.jsp>.

Heinrichs W., Srivastava S., Montgomery K., and Dev P. The fundamental manipulations of surgery: a structured vocabulary for designing surgical curricula and simulators *J Am Assoc Gynecol Laparosc* 2004; 4 (11): 450-6.

HL7: <http://www.hl7.org/>. [Last accessed December 5, 2005.]

Howell, K. and Higgins, G.. *Education Framework for Medical Simulation Training*. Federation of American Scientists. August 2004.
<http://www.fas.org/main/content.jsp?formAction=297&contentId=271>. [Last accessed March 2006].

Hubal, R., Kizakevich, P., Guinn, C., Merino, K., and West, S. (2000), "The Virtual Standardized Patient – Simulated Patient Practitioner Dialogue for Patient Interview Training" in J.D. Westwood, H.M., Hoffman, G.T. Mogel, R.A. Robb, and D.

Stredney (Eds.), *Envisioning Healing: Interactive Technology and the Patient Practitioner Dialogue*. IOS Press: Amsterdam.

Hubal, R., Kizakevich, P., Guinn, C., Merino, K., and West, S. (2000). The Virtual Standardized Patient: Simulated Patient-Practitioner Dialogue for Patient Interview Training. In J.D. Westwood, H.M. Hoffman, G.T. Mogel, R.A. Robb, and D. Stredney (Eds.).

Issenberg, B., Mcgaghie, W., Petrusa, E., Lee, G., Scalese, R. (2005). Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Medical Teacher* (ISSN: 0142-159X); Volume 27, No. 1, pp. 10-28(19); January.

Jacobs, J.W., C. Prince, R.T. Hays and E. Salas (1990) A meta-analysis of the flight simulator training research. Human Factors Division, Naval Training Systems Center. NAVTRASYSCEN TR-89-006.

Jorna, P.G., E.R. Van Kleef and W.P. de Boer (1992) Aircraft simulation and pilot proficiency: From surrogate flying towards effective training. National Aerospace Laboratory, Amsterdam, Netherlands. NLR-TP-91489-U, ADB179 300.

Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development*, 48, 63-85..

Kanfer, R. and McCombs, B. L. (2000). Motivation: Applying current theory to critical issues in training. In S. Tobias and J. D. Fletcher (Eds.), *Training and retraining: A handbook for business, industry, government, and the military* (pp. 85-108). New York, NY: Macmillan.

Knerr, C.M., J.E. Morrison, R.J. Mumaw, D.J. Stein, P.J. Sticha, R.G. Hoffman, D.M. Buede and D.H. Holding (1986) *Simulation-Based Research in Part-Task Training*. Air Force Human Resources Laboratory, DTIC AD-B107293.

Koedinger, K. R., Anderson, J. R., Hadley, W. H., and Mark, M. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education*, 8, 30-43.

Kirkpatrick, D. L. (1959, 1960). Techniques for evaluation training programs. *Journal of American Society for Training Directors*, 13, 28_32.

Kiss, B., G. Szijártó, B. Benedek, L. Simon, G. Csukly, B. Takács (2003), "CyberTherapy: Applications of Virtual Reality and Digital Humans in Clinical Psychology", in 2nd Intl. Conf. On Computer Animation & Geometric Modeling, Budapest, Hungary.

Kurtz S, Silverman J, Draper J. *Teaching and learning communication skills in medicine*. Oxford: Radcliffe Medical Press, 1998.

MedTech. *U.S. Market for Critical Care Patient Management Products* (2004).

•
•
•
•
•
•
•
•

Lee, A.T. and S.R. Bussolari (1989) Flight simulator platform motion and air transport pilot training. *Aviation, Space and Environmental Medicine* 60:136-140.

Lintern, G. and S. Roscoe (1978) Transfer of landing skill after training with supplementary visual cues. USAF Academy's Sixth Symposium on Psychology in the Department of Defense.

Lintern, G. (1980) Transfer of Landing Skill After Training with Supplementary Visual Cues. *Human Factors* 22(1):81-88.

Lintern, G. and R.S. Kennedy (1984) Video game as a covariate for carrier landing research. *Perceptual and Motor Skills* 58:167-172.

Lintern, G., K.E. Thomley-Yates, B.E. Nelson and S.N. Roscoe (1987) Content, variety, and augmentation of simulated visual scenes for teaching air-to-ground attack. *Human Factors* 29:45-59.

Lintern, G., D.J. Shepard, D.L. Parker, K.E. Yates and M.D. Nolan (1989) Simulator design and instructional features for air-to-ground attack: A transfer study. *Human Factors* 31:87-99.

Lintern, G., S.N. Roscoe and J.E. Sivier (1990) Display principles, control dynamics and environmental factors in pilot training and transfer. *Human Factors* 32:299-317.

Lintern, G., S.N. Roscoe, J.M. Koonce and L.D. Segal (1990) Transfer of landing skills in beginning flight training. *Human Factors* 32:319-327.

Lintern, G., H.L. Taylor, J.M. Koonce, R.H. Kaiser and G.A. Morrison (1997) Transfer and quasi-transfer effects of scene detail and visual augmentation in landing training. *The International Journal of Aviation Psychology* 7:149-169.

Loesch, R.L. and J. Waddell (1979) The importance of stability and control fidelity in simulation. *Proceedings, 50 Years of Flight Simulation*. London. Royal Aeronautical Society. Pp 90-94.

Marcus, A.J. and L.E. Curran (1988) The use of flight simulators in measuring and improving training effectiveness. Center for Naval Analyses, Naval Planning Manpower and Logistics Division. CRM-86-27, AD A171 661.

Martin, E.L. (1981) Training effectiveness of platform motion: Review of motion research involving the advanced simulator for pilot training and the simulator for air-to-air combat. Operations Training Division, Air Force Human Resources Laboratory. AFHRL-TR-79-51, ADA 095 930.

Martin, E. and W. Waag (1978a) Contributions of platform motion to simulator training effectiveness: Study I-Basic contact. Air Force Human resources Laboratory. AFHRL-TR-78-15.

Martin, E. and W. Waag (1978b) Contributions of platform motion to simulator training effectiveness: Study I - Aerobatics. Air Force Human resources Laboratory. AFHRL-TR-78-52.

Martin, E.L. and D.F. Cataneo (1980) Computer generated image: Relative training effectiveness of day versus night visual scenes. Air Force Human resources Laboratory. AFHRL-TR-79-56.

Mayer, R. E. (1999). Designing instruction for constructivist learning. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: A new paradigm of instructional theory*, Vol. II. (pp. 141-159). Mahwah, NJ: Lawrence Erlbaum Associates.

Mayer, R. E. (2001). *Multimedia learning*. Cambridge, England: Cambridge University Press.

McCauley, M.E. and J.C. Cotton (1982) VTRS carrier landing, transfer of training, and performance measurement development. Canyon Research Group.

Mehta, N., Haluck, R., Frecker, M., Snyder, A., (2002). Sequence and task analysis of instrument use in common laparoscopic procedures. *Surg Endosc* 16:280-285.

MedBiquitous: http://meld.medbiq.org/primers/virtual_patients_cases_ellaway.htm. [Last accessed December 5, 2005.]

Merrill, M. D. (1983). Component display theory. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: An overview of their current status* (pp. 279-333). Hillsdale, NJ: Lawrence Erlbaum Associates.

Merrill, M. D. (1999). Instructional transaction theory (ITT): Instructional design based on knowledge objects. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: A new paradigm of instructional theory*, Vol. II. (pp. 397-424). Mahwah, NJ: Lawrence Erlbaum Associates.

Merrill, M. D. (2002). A pebble-in-the pond: Model for instructional design. *Performance Improvement*, 41, 39-44.

Micheli, G. (1972) Analysis of the transfer of training, substitution and fidelity of simulation of transfer equipment. Naval Training Equipment Center. TAEG Report 2.

Montemerlo, M. (1977) Training device design: The simulation/stimulation controversy. Naval Training Equipment Center. NAVTRAEQUIPCEN IH-287.

Morgan PJ, Cleave-Hogg D, DeSousa S, et al. High-fidelity patient simulation: validation of performance checklists. *Br J Anaesth* 2004;92:388-92.

•
•
•
•
•
•
•
•

Moroney, W.F. and B.W. Moroney (1998) Flight simulation, In: Handbook of Aviation Human Factors, D.J. Garland, J.A. Wise, V.D. Hopkin (Eds.). New York, Erlbaum.

Nackman GB, Bermann M, Hammond J. Effective use of human simulators in surgical education. J Surg Res 2003;115:214–18.

National Research Council. 2001. *Knowing what students know: The science and design of educational assessment*. Committee on the Foundations of Assessment. Pelligrino, J., Chudowsky, N., and Glaser, R., editors. Board on Testing and Assessment, Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Nataupsky, M., W. Waag, D. Weyer, R. McFaddeb and E. McDowell (1979) Platform motion contribution to simulator training effectiveness: Study III: Interaction of motion with field of view. Air Force Human Resources Labortaory. AFHRL-TR-79-25.

Norcini J , Boulet J. Methodological issues in the use of standardized patients for assessment. Teach Learn Med 2003;15:293–7.

Orlansky, J., C.J. Dahlman, C.P. Hammon, J. Metzko, H.L. Taylor and C. Youngblut (1994) The Value of Simulation for Training. Institute for Defense Analysis, Report IDA P-2982.

Orlansky, J., H.L. Taylor, D.B. Levine and J.G. Honig (1997) Cost and Effectiveness of the Multi-Service Distributed Training Testbed (MDT2) for Training Close Air Support. Institute for Defense Analysis, NTIS ADA327227.

Orlansky, J. and String, J. (1979). Cost-Effectiveness of Computer-Based Instruction in Military Training. Alexandria, VA: Institute for Defense Analyses, Report IDA P-1375.

Perkins, D. N. and Unger, C. (1999). Teaching and learning for understanding. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: A new paradigm of instructional theory*, Vol. II. (pp. 91-114). Mahwah, NJ: Lawrence Erlbaum Associates

Pfeiffer, M. and W. Rankin (1985) Training effectiveness evaluation of passive acoustic analysis trainer. Naval Training Systems Center. TR-86-019.

Pfeiffer (1983) Experimental and analytic evaluation of the effects of visual and motion simulation in SH-3 helicopter training. TR-85-002.

Pfeiffer, M.G., J.D. Horey and S.K. Butrimas (1991) Transfer of simulated instrument training to instrument and contact flight. *The International Journal of Aviation Psychology* 1:219-229.

Pham T, Roland L, Benson KA, Webster RW, Gallagher AG, Haluck RS. Smart tutor: a pilot study of a novel adaptive simulation environment. *Stud Health Technol Inform.* 2005;111:385-9.

Pozner, C., Spilseth, E., Barker, T, Takayesu, J, Nadel, E., Brown, C., Bhatia, K., Gatewood, M., Look R., Nelson S., Walls, R. Simulation as an Integral Component of an Emergency Medicine Residency at Harvard.
http://www.anestech.org/media/Publications/IMMS_2005/Ponzer.pdf

Ramachandran S., Cramer M., Harville D., Ashworth A. Developing an Adaptive ADL Solution for Training Medical Teams. Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2003.

Reiser, B., Tabak, I, Sandoval, W. A., Smith, B. K., Steinmuller, F., and Leone, A. J. (2001). BGuILE: Strategic and conceptual scaffolds for scientific inquiry in biology classrooms. In S. M. Carver and D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 263-305). Mahwah, NJ: Lawrence Erlbaum Associates.

Rizzo, A.A. (2002), Virtual Reality and Disability: Emergence and Challenge” in *Disability and Rehabilitation*, vol. 24, no. 11, 2002, pp. 567-569.

Ruttkay Z., Pelachaud, C. (Eds.), *From Brows to Trust: Evaluating Embodied Conversational Agents*, Kluwer, 2004.

Rankin, W., P. Bolton, R. Shikiar and L. Saari (1984) Simulator fidelity and training effectiveness: A comprehensive bibliography with selected annotations. U.S. Nuclear Regulatory Commission. NUREG/CR-3726.

Suthers, D., Weiner, A., Connelly, J., and Paolucci, M. (1995). Belvedere: Engaging students in critical discussion of science and public policy issues. In *Proceedings of AIED-95*.

Salas, E. and Cannon-Bowers, J. A. (1997). Methods, tools, and strategies for team training. In M. A. Quinones and A. Ehrenstein (Eds.), *Training for a rapidly changing workplace: Applications of psychological research* (pp. 249-279). Washington, D. C.: American Psychological Association.

Salas, E. and Cannon-Bowers, J. A. (2000). The anatomy of team training. In S. Tobias and J. D. Fletcher (Eds.), *Training and retraining: A handbook for business, industry, government, and the military* (pp. 312-335). New York, NY: Macmillan.

Satava, RM, (1996) Cybersurgeon: Advanced simulation technologies for surgical education”. *Bulletin of the American College of Surgeons*, July: 77-81.

Schank, R. C., Berman, T. R., and Macpherson, K. A. (1999). Learning by doing. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: A new paradigm of instructional theory*, Vol. II. (pp. 161-181). Mahwah, NJ: Lawrence Erlbaum Associates.

•
•
•
•
•
•
•
•

Schaefer, J.J., Dongelli, T., Gonzalez, R.M. (1998) Results of systematic psychomotor difficult airway training of residents using the ASA difficult airway algorithm and dynamic simulation, *Anesthesiology*, 89(3A), Supplement, A60.

Schraagen, J., Chipman, S., Shalin, V., (eds). *Cognitive Task Analysis* (2000). Lawrence Erlbaum Publishing, MahWah, NJ.

Seymour, Neal E. MD *; Gallagher, Anthony G. PhD +; Roman, Sanziana A. MD *; O'Brien, Michael K. MD *; Bansal, Vipin K. MD *; Andersen, Dana K. MD *; Satava, Richard M. MD. *Virtual Reality Training Improves Operating Room*

Performance: Results of a Randomized, Double-Blinded Study. (2002) *Annals of Surgery*. 236(4):458-464, October.

Spears, W.D. (1983) *Processes of skill performance: A foundation for the design and use of training equipment.* Naval Equipment Training Center, DTIC ADA13679.

Spunt, D., Schaivone, K., *Raising the Bar: Improving Clinical Competency Assessment through Clinical Simulation and Standardized Patients.*

mtc-europe.co.uk/Publications/IMMS_2004/Spunt.pdf.

Sterling, B.S. (1993a) *Impact of tank PGT training on successive Tank Table XII scores.* Grafenwoehr, Germany, Seventh Army Training Command. Training Note #7.

Sterling, B.S. (1993b) *Use of the M1 Platoon Gunnery Trainer to Train Battle Command Skills.* Grafenwoehr, Germany, Seventh Army Training Command. Training Note #11.

Sterling, B.S. (1994) *Use of the Bradley Platoon Gunnery Trainer to Train Battle Command Skills.* Grafenwoehr, Germany, Seventh Army Training Command. Training Note #17.

Sterling, B.S. (1996) *Relationship Between Platoon Gunnery Training and Live-Fire Performance.* U.S. Army Research Institute for the Behavioral and Social Sciences, NTIS ADA319342.

Tannenbaum, S. I., Mathieu, J. E., Salas, E., Cannon-Bowers, J. A. (1991). *Meeting trainees' expectations: The influence of training fulfillment on the development of commitment, self- efficacy, and motivation.* *Journal of Applied Psychology*, 76, 759-769.

Tannenbaum, S. I. and Yukl, G. (1992). *Training and development in work organizations.* *Annual Review of Psychology*, 43, 399-441.

Thomas, G.S., M.R. Houck and H.H. Bell (1990) Training evaluation of air combat simulation. Operations Training Division, Air Force Human resources Laboratory. AFHRL-TR-90-30, B145631L.

U.S. Markets for Critical Care Patient Management Products (2004). MedTech Insight Market Research Report, A340.

van Merriënboer, J. J. G. (1997). Training Complex Cognitive Skills. Englewood Cliffs: Educational Technology Publications.

Voss, H., W. Boney and W. Rankin (1970) Prototype instinctive firing training device for small arms. Naval Training Device Center. NAVTRADEVCEEN IH-179.

Westra, D.P., C.W. Simon, S.C. Collyer and W.S. Chambers (1981) Simulator Design Features for Carrier Landing. I. Performance Experiments, Naval Training Equipment Center, Technical Report 78-C-0060-7.

Westra, D.P. (1982) Simulator Design Features for Carrier Landing: II. In-Simulator Transfer of Training. Naval Training Equipment Center, DTIC 81-C-0105-1.

Westra, D.P., G. Lintern, D.J. Sheppard, K.E. Thomley, R. Mauk, D.C. Wightman and W.S. Chambers (1988) Simulator Design and Instructional Features for Carrier Landing: A Field Transfer Study. Naval Training Systems Center, Technical Report 85-C-0044-2.

Youngblood P., Srivastava S., Curet M, Heinrichs W., Dev P., Wren S. Comparison of training on two laparoscopic simulators and assessment of skills transfer to surgical performance. *J Am Coll Surg* 2005; 4 (200): 546-51.

<http://bio.ltsn.ac.uk/journal/vol3/beej-3-7.htm>. [Last access 5 December 2005].

•
•
•
•
•
•
•



VIRTUAL PATIENT WORKSHOP PARTICIPANTS [as of June, 2005]

- Paul Barach, MD. MPH, Department of Anesthesiology, University of Miami
- Randy Brown, DVM, Associate Director, Johns Hopkins Simulation Center,
Assistant Professor Johns Hopkins
- Beth Bryant, Digital Mill
- Adam Burrowbridge, Researcher Federation of American Scientists
- Jan Cannon-Bowers, Ph.D., Associate Professor, School of Film, University of
Central Florida, School of Film
- Dexter Fletcher, Ph.D., Institute for Defense Analyses
- Dawn Foster, RN, MS, CCRN University of Maryland
- Robert Foster, Ph.D. Program Director, Office of the Secretary of Defense (DDR&E)
- Anthony Gallagher, Ph.D. Emory University
- Gerry Higgins, Ph.D., Laerdal Medical Corporation, Washington, D.C.
- Kay Howell, Federation of American Scientists
- Jeff Jacoby, Medical Education Technologies, Inc
- Bruce Jarrell, MD, FACS Academic Dean, University of Maryland School of
Medicine
- Christoph Kaufmann, MD, MPH, Portland Trauma Service; Advanced Trauma Life
Support Subcommittee, American College of Surgeons
- Henry Kelly, Ph.D., Federation of American Scientists
- Andrew Kofke, MD, MBA, FCCM Professor, Director of Neuroanesthesia
Departments of Anesthesia and Neurosurgery, University of Pennsylvania
- Fred Kron, MD Assistant Clinical Professor, Department of Family Medicine,
University of Wisconsin-Madison
- Tore Laedal, President and CEO, Laerdal Medical
- Michelle Lucey-Roper, D.Phil Federation of American Scientists
- Harvey Magee, Telemedicine and Advanced Technology Research Center (TATRC)

Mary Beth Mancini, RD, Ph.D., CAN, FAAN Professor, Associate Dean for Undergraduate Nursing Programs, University of Texas at Arlington

Lance Manning, Prenosis

Rudy McDaniel, Ph.D., University of Central Florida

Gil Muniz, Ph.D., National Capital Area Simulation Center, Uniformed Services University

Vinay Nadkarni, MD; Department: Anesthesiology; Division: Critical Care Medicine, Children's Hospital of Philadelphia

Sergei Nirenburg, Ph.D.

Amar Patel University of Maryland

Sachin Patil, Federation of American Scientists

Clive Patrickson, Ph.D., MBA, JD, Laerdal Medical

Christine Pintz, RNC, MSN, FNP University of Maryland

John Raczek, University of Maryland

Daniel Raemer, Ph.D. Center for Medical Simulation, Harvard (Massachusetts General Hospital)

Sowmya Ramachandran, Ph.D. Stottler-Henke Associates

Cheryl Robertson, RN, MSN, APRN, BC (ANP, WHCNP-C University of Maryland

Ben Sawyer Digitalmill, Serious Initiative

Ross Scalese, MD, FACP, Assistant Professor of Medicine, Assistant Director, Educational Research and Technology -Center for Research in Medical Education, Department of Medicine, University of Miami

Marc Scerbo, Ph.D. Old Dominion University

John Schaefer, MD, Director, Wisser, UPMC

Kathy Schaivone, MPA, Clinical Education and Evaluation Lab University of Maryland

Mark Schleicher, Federation of American Scientists

Steve Schmidt, SimMedical

•
•
•
•
•
•
•
•

Steve Senager, Ph.D.

Azizeh Sowan RN, MS University of Maryland

Debra Spunt, MS, RN Director, Clinical Simulation, University of Maryland School of Nursing

Walt Stoy, Ph.D., EMT-P, CCEMTP Professor and Director of Emergency Medicine, University of Pittsburgh; Project Director for EMT-I and Paramedic, for all the National Standard Curriculum.

Ron Walls, MD, Chairman, Department of Emergency Medicine, Harvard Medical School

Bruce Walz, Ph.D., NREMT-P, Associate Professor and Chairmen Department of Emergency Health Services, University of Maryland Baltimore County

Doug Whatley CEO, Breakway Games

Erling Woods, Ph.D., Global Vice-President, Research and Development Laerdal Medical

Patricia Youngblood, MA, M. Ed, Ph.D., Associate Director, Evaluation Stanford University School of Medicine

Virtual Patient Workshop – Final Agenda
University of Maryland School of Medicine

Monday, June 27th 2005

8:15 am: - "Introduction and Objectives – What is the Research Roadmap?" - Henry Kelly, Ph.D., Federation of American Scientists - The Learning Federation

8:40 am - "Computerized Patient Simulation for Learning Cognitive and Psychomotor Skills,"
- Gerry Higgins, Ph.D., Laerdal Medical Corporation

9:00 am – PANEL: Computer-based Training in EMS and Nursing Education
Moderator: Gerry Higgins, Ph.D.

- "Simulation and EMS: The edge of possibility..." - Walt Stoy, Ph.D., EMT-P, CCEMTP, Professor and Program Director, Center for Emergency Medicine, University of Pittsburgh
- "Clinical Simulation Labs: The Golden Thread which Connects a Curriculum," - Debra Spunt, M.S., R.N., University of Maryland School of Nursing

10:00 am - *BREAK with Coffee, Tea, Fruit, Water and Pastries*

10:15 am – PANEL: The Future of Learning Technologies is Here Now
Moderator: Kay Howell

- "Automated Tutoring for Simulation-based Training" – Sowmya Ramachandran, Ph.D. Stottler-Henke Associates
- "Scenario-based Immersive Training for High Performance Teams" - Jan Cannon-Bowers, Ph.D., Institute for Simulation and Training, University of Central Florida

11:30 am – PANEL: Current Status and Vision from the Medical Simulation Centers

Moderator: Bruce E. Jarrell, M.D., F.A.C.S.

- "Simulation: A Surgeon's Cautious Optimism" - Christoph Kaufmann, M.D., M.P.H., F.A.C.S., Associate Medical Director, Trauma Services, Legacy Emanuel Hospital, Portland; Chair, Committee on Advanced Trauma Life Support (ATLS), American College of Surgeons; Former Director, Surgical Simulation Lab, USUHS National Capital Area Simulation Ctr.; Professor of Surgery, Uniformed Services University and Clinical Associate Professor of Surgery, Oregon Health Sciences Center



- “Competency-based Simulation Training in Anesthesiology at the University of Pittsburgh ” - John J. Schaefer, III, M.D., Director, Peter M. Winter Institute for Simulation, Education and Research, University of Pittsburgh
- “Automated Virtual Human from Medical Knowledge” - Bruce E. Jarrell, M.D., F.A.C.S., Vice Dean for Academic Affairs, University of Maryland School of Medicine.
- “Would you let me operate on your uncle?” - Daniel Raemer, Ph.D. Harvard Simulation Center; Director, Board, Society for Medical Simulation

12:45 pm – *BREAK* followed by **WORKING LUNCH**

2:00 pm - “Charge to Working Groups”

2:15 pm – **WORKING GROUPS** – Define initial research roadmaps:

1. Intelligent Generation of Patient Case Scenarios in Medical Simulation
2. Technology Systems for Learner Modeling with an Emphasis on Intelligent Coaching in the Simulation Environment
3. Debriefing and Assessment in Medical Simulation-based Training

4:45 pm - Working Group #1 – Brief review and revise output from other working groups

4:45 pm - Working Group #2 – Brief review and revise output from other working groups

4:45 pm - Working Group #3 – Brief review and revise output from other working groups



Tuesday, June 28th 2005

9:00 am – **Review of activities and mission:** Kay Howell, Gerry Higgins, Ph.D., Bruce E. Jarrell, M.D., F.A.C.S.

- Where do we stand? What are the Issues? How do we author the Research Roadmap?
- When and where will the Research Roadmap be published? [e.g., *Academic Medicine?*]

9:45 am - PANEL: Breakthroughs and Opportunities in Medical Simulation

Moderator: Gerry Higgins, Ph.D.

- "Virtual Reality Training for the Operating Room and Cath. Lab: A Paradigm Shift in Training for Procedural Based Medicine" - Tony Gallagher, Ph.D., Emory University
- "STRATUS Center for Medical Simulation: Vision for the next 10 years" - Ron Walls, M.D., Harvard Medical School
- "Simulation for Multidisciplinary Mock Codes" - Elizabeth A. Hunt, M.D., MPH, Director, Johns Hopkins Simulation Center, Johns Hopkins School of Medicine, Department of Anesthesiology and Critical Care Medicine

11:15 am – Presentations from Each of the Working Group Leaders, and Discussion

1:00 pm – WORKING LUNCH (*continuing discussion*)

2:00 pm – Wrap-up and Discussion of Next Steps, Kay Howell, Gerry Higgins, Ph.D.