MILITARY MEDICINE, 178, 10:76, 2013

Prevention of Surgical Skill Decay

Ray S. Perez, PhD; Anna Skinner, MA†; Peter Weyhrauch, PhD‡; James Niehaus, PhD‡; Corinna Lathan, PhD, PE†; Steven D. Schwaitzberg, MD, FACS§; Caroline G. L. Cao, PhD‖*

ABSTRACT The U.S. military medical community spends a great deal of time and resources training its personnel to provide them with the knowledge and skills necessary to perform life-saving tasks, both on the battlefield and at home. However, personnel may fail to retain specialized knowledge and skills if they are not applied during the typical periods of nonuse within the military deployment cycle, and retention of critical knowledge and skills is crucial to the successful care of warfighters. For example, we researched the skill and knowledge loss associated with specialized surgical skills such as those required to perform laparoscopic surgery (LS) procedures. These skills are subject to decay when military surgeons perform combat casualty care during their deployment instead of LS. This article describes our preliminary research identifying critical LS skills, as well as their acquisition and decay rates. It introduces models that identify critical skills related to laparoscopy, and proposes objective metrics for measuring these critical skills. This research will provide insight into best practices for (1) training skills that are durable and resistant to skill decay, (2) assessing these skills over time, and (3) introducing effective refresher training at appropriate intervals to maintain skill proficiency.

INTRODUCTION

Prevention of skill decay is and should be a high priority for any training organization. Skill decay is the partial or full loss of trained or acquired skills and knowledge following periods of nonuse.[1] Loss of critical skills and knowledge is of great concern, and is problematic in situations in which medical personnel receive initial training but may not have an opportunity to use these skills or knowledge for extended periods of time. Arthur et al[1] in a meta-analysis examining the results of 53 articles within the skill decay and retention literature, found that after 365 days of nonuse or nonpractice, the average participant's performance was reduced by almost a full standard deviation ($d = -0.92$). There is evidence that declarative knowledge (e.g., facts, principles, and concepts) decays at a slower rate than procedural knowledge (e.g., multiple steps that must be performed in a specified order to solve a problem or complete a task).[2] Psychomotor skills also appear to be more resistant to skill decay than cognitive tasks, in the absence of mental rehearsal.[1] Although the research literature on skill decay is comprehensive and extensive, this is not true of the literature specifically pertaining to laparoscopic surgical (LS) skill decay. One study that did focus on these skills was conducted by Brunner and Korndorffer[3]; they examined skill decay of LS skills, echoing the findings by Arthur and his colleagues, demonstrating that LS skills could be trained in a virtual reality learning environment and that nonuse exhibited the same pattern of skill decay reported in the meta-analysis conducted by Arthur et al[1] Despite this specific conclusion, a primary limitation within this area is the lack of research examining the acquisition and retention of the cognitive skills of LS (e.g., decision making and problem solving). Within the domain of general surgery, Jacklin et al[4] studied the effects of providing cognitive feedback to surgical trainees in an attempt to improve their risk assessment of surgery (i.e., judgment of postoperative mortality risk). They found that the use of cognitive feedback improved the accuracy of trainees' estimates of postoperative mortality risk. However, very little research has been conducted to examine the skill decay of these critical cognitive skills, and similar research has not been conducted within the LS domain.

Section 1 of this article reviews skill acquisition and retention research related to general skills, military skills, and LS skills. Section 2 introduces models that identify critical skills related to laparoscopy, and Section 3 proposes objective metrics for measuring these critical skills. Section 4 addresses the development of skill decay curves, and Section 5 proposes a model that examines refresher training of laparoscopic skills. The article concludes with a summary and conclusions that include future areas of research. This article includes two perspectives and approaches to addressing the issues surrounding LS skill assessment, training, and sustainment: (1) the development and use of a Laparoscopic Surgery Training System (LASTS) prototype for learning, refreshing, and assessing LS skills and (2) the design, development, and validation of a portable, open architecture Surgical Skills Training and Assessment Instrument (SUSTAIN) to support acquisition and retention of fundamental psychomotor, perceptual, and cognitive laparoscopic skills. Both systems seek to leverage the development and validation of an empirical model of skill acquisition and decay.

*Office of Naval Research, 875 North Randolph Street, Arlington, VA 22203.

†AnthroTronix, Inc., 8737 Colesville Road, L203 Silver Spring, MD 20910.

‡Charles River Analytics, Inc., 625 Mount Auburn Street, Cambridge, MA 02138.

§Cambridge Health Alliance, 1493 Cambridge Street, Cambridge, MA 02139.

‖ Biomedical, Industrial and Human Factors Engineering, Wright State University, 3640 Colonel Glenn Hwy, Dayton, OH 45435.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of any of the funding agencies.

doi: 10.7205/MILMED-D-13-00216

SECTION 1: BACKGROUND

Skill Acquisition and Decay

To analyze skill decay, we must first examine the process by which knowledge and skills are acquired. The rate of skill acquisition depends on many factors; however, there is evidence that rates of learning tend to follow a logarithmic or exponential function, rather than a linear relationship between the time to perform a task and the number of practice attempts.[5,6] There are typically three stages of skill development: a cognitive stage, associative stage, and an autonomous stage; it is during the autonomous stage that expertise is achieved.[7] Research has shown that the major factors influencing skill decay and retention include the length of retention (nonuse) interval, degree of overlearning, task characteristics (e.g., closed loop versus open loop, physical versus cognitive), methods of testing for initial learning and retention, conditions of retrieval (recognition versus recall), instructional strategies or training methods, and individual differences in trainees.[2]

Skill Acquisition and Decay Within Military Tasks

The rate of skill decay within military tasks has been researched extensively in the past. For example, Wisher et al[2] investigated the decay of skills and knowledge with 20,000 reservists, and found that gross motor skills decayed after approximately 10 months, whereas cognitive skills, such as recall of procedures, decayed within approximately 6 months. Thus, while both motor and cognitive skills are subject to decay with nonuse, the respective rates of decay differ. The Naval Education and Training Command developed a categorization of navy tasks on a scale of proneness to decay; tasks most prone to decay included recalling procedures and voice communications tasks, whereas those least prone to decay included gross motor skills and attitude learning.

Wisher et al[2] also conducted an extensive review of general skill acquisition and retention/decay literature. Based on this review, they categorized military tasks into 3 components: (1) knowledge, (2) decision, and (3) execution. The knowledge category is based on the recall of domain-specific information; the decision category depends on cognitive processing of the domain-specific information; and the execution category refers to both the perceptual and motor requirements of a task. Wisher et al[2] also identified specific task factors that affect skill acquisition and decay, such as task complexity and task demands. Other factors identified as affecting decay included task time pressure, whether or not job aids were used (job aids decreased decay), and the quality of job aids used (higher quality job aids decreased decay).

Although past research efforts have provided a greater understanding of the mechanisms underlying skill acquisition and decay, including the relative rates of decay of various types of skills, there are no detailed models and skill decay curves of military medical tasks. A need exists for validated models and skill decay curves within the context of specialized military medical skills, such as LS skills, because these specialized procedures are not usually performed during deployments, and are therefore susceptible to decay. Validated LS skill decay models would support the development of guidelines for accurately timed refresher training to prevent LS skills decay.

Skill Acquisition and Decay Within Laparoscopic Surgical Tasks

Over the past decade, attempts have been made to establish standards for laparoscopic surgical skills training and evaluation. The McGill Inanimate System for Training and Evaluation of Laparoscopic Skills (MISTELS) has been shown to be reliable and valid as an educational tool.[8] This method of assessment has been incorporated into the manual skills training practicum portion of the Society of American Gastrointestinal and Endoscopic Surgeons' (SAGES) Fundamentals of Laparoscopic Surgery (FLS) training program, and includes a portable video trainer box for rehearsal of five basic manual skills tasks: peg transfer, pattern cutting, endo-loop placement, extracorporeal suturing, and intracorporeal suturing (see Fig. 1).

FLS trainer scores have been shown to be predictive of intraoperative laparoscopic performance as measured by the Global Operative Assessment of Laparoscopic Skills (GOALS) manual assessment framework,[9] making the FLS training program the current "gold standard" in laparoscopic training, and resulting in its rapid adoption as a primary component of many general surgery residency programs.[10,11]

FLS manual skills training results in laparoscopic manual skills that have shown to be durable for up to 11 months,[12] and skill retention has been demonstrated for three tasks similar to FLS tasks for up to one year.[13] However, these studies relied only on time to complete the specified manual skills tasks, and retention of the cognitive components of LS training was not considered. Also, although maintenance of laparoscopic manual skills through rehearsal and retraining has been shown to prevent decay,[12] no standards currently exist for retraining, and few deployable systems exist that can be used where they are most needed—in far forward military

FIGURE 1. FLS box trainer.

medical facilities—to provide critical skills refresher training during long deployments.

Stefanidis et al[14] provided preliminary evidence that a visual–spatial secondary task assessing spare attentional capacity may help distinguish among individuals of variable laparoscopic expertise when standard FLS performance measures fail to do so, and that automaticity metrics may improve current simulator training and assessment methods. Stefanidis et al[15] tested this theory further, demonstrating that after approximately 10 hours of training (average of 84 trials) on the FLS intracorporeal suturing task with concurrent performance of a visual–spatial secondary task, novices demonstrated improvements in both suturing and secondary-task performance compared with baseline scores; however, none of the subjects achieved expert level secondary task proficiency. Recently, Stefanidis et al[11] demonstrated that training to automaticity (overlearning) on the FLS intracorporeal suturing task while concurrently performing a visual–spatial secondary task resulted in improved operating room performance on a porcine model as compared to training to proficiency without the secondary task. This level of training required, on average, 163 trials, and nearly half of the participants assigned to the automaticity training group were unable to achieve expert performance levels on the secondary task. Therefore, while superior training and skill transfer was demonstrated, the training costs are high in terms of time and resources. Also, the secondary task used was unrelated to surgical tasks. It is possible that by incorporating a secondary task that addresses interoperative skills, including cognitive skills, training would be enhanced further, providing justification for the extended time associated with training to automaticity. This article not only explores the impact of automaticity on the acquisition and durability of an LS skill but also raises a number of questions, such as: (1) How much training is actually necessary to competently perform LS?; (2) Do fundamental LS skills that become automated generalize to other minimally invasive surgery skills and specific procedures?; and (3) What is the impact of automaticity on the training of cognitive LS skills?

Recent research has sought to develop a deeper understanding of the cognitive factors involved in training. For example, Park et al[16] have attempted to understand the role of declarative memory processes during psychomotor skills learning using a cognitive simulation based on the Adaptive Control of Thought—Rational model. However, they did not directly assess the role of decision making involved in surgical procedures. Palter,[17] in a review of training curriculum for Minimally Invasive Surgery (MIS), suggests that MIS training should include cognitive teaching to compensate for knowledge gaps.

Finally, given that LS tasks require surgeons to rely heavily on 2-dimensional visual cues and haptic cues to perform complex tasks within a 3-dimensional space, visual–spatial skills are likely to play an integral role in the performance of these tasks. Visual–spatial skills involve the perception and processing of spatial relationships, such as mental transformation (e.g., mental rotation) and relative distance perception. Hassan et al[18] found that among novices, visual–spatial perception is associated with manual skills performed on a laparoscopic skills virtual reality (VR) simulator; novice participants with a high degree of spatial perception performed laparoscopic VR tasks faster than those who had a low degree of spatial perception and also scored better for economy of motion, tissue damage, and total error. Ritter et al[19] demonstrated that spatial perceptual abilities correlated well with duration of the learning curve (i.e., number of trials required to meet the specified proficiency criterion) on a validated VR flexible endoscopy simulator. Visual–spatial ability assessments included determination of the 3-dimensional orientation of a 2-dimensional grayscale cube via the pictorial surface orientation (PicSOr) test and the Cube Comparison test, which involves mental rotation of an object about its center. The PicSOr test was also shown in three studies by Gallagher et al[20] to consistently predict performance on the FLS circle cutting task, as well as significantly predicting laparoscopic surgeons' performance. Keehner et al[21] also examined changes in performance as novices learned to use an angled laparoscope within a virtual environment; initial performance showed considerable variability among novices, with performance related to both general, nonverbal reasoning ability (assessed via the abstract reasoning task of the Differential Aptitude Test battery) and spatial abilities (assessed using the Mental Rotation Test and Visualization of Views Test). As learning progressed, the correlation of performance with general reasoning ability diminished after the first few sessions, whereas by contrast, the significant correlation with spatial ability persisted even after the group variance had diminished. This finding provided further support for a previous study by Keehner et al,[22] which found a significant correlation between spatial ability and interoperative videoscopic skills performed on animals for a novice group; however, no significant correlation was found within a group of experienced surgeons. Thus, the importance of spatial ability in performance of laparoscopic skills seems to diminish with experience.

SECTION 2: CRITICAL SKILLS

The first step in creating a system to train and refresh LS skills is to identify and model the critical underlying skills. Currently, no standard method exists for identifying critical skills related to laparoscopy. Current understanding is also limited on how best to refresh decayed skills, and no standard refresher training currently exists. Previous task analyses of (MIS) procedures have focused on psychomotor skills.[23] However, the psychomotor ability to operate the laparoscopic tools represents only a portion of the skills needed to successfully perform MIS. Visual–spatial skills are needed to orient within the abdomen and recognize structures such as tissue, organs, ducts, and blood vessels. Cognitive skill is needed to maintain proper situational awareness, understand the ramifications of observations, make the correct decisions, and act in the best interest of the patient and the operation. In addition, different surgeons perform procedures with different techniques, based on varying degrees of expertise and experience, as well as training background. For example, establishing initial laparoscope and instrument access during LS cholecystectomy can be performed using the Hassan technique or using the Veress needle technique, and expert opinions differ on which is optimal. Therefore, the model of skills must account for both normative errors and actions that are not correct in any of these techniques, as well as quasinormative errors, which consist of actions that are correct for some techniques but not others.[24] The following details two complementary approaches to identifying and modeling the critical skills underlying LS proficiency, with the goal of developing improved training, assessment, and maintenance of these skills.

LASTS Critical Skills

The LASTS uses models of surgical skill acquisition and decay, objective methods of surgical skill assessment, and individual skill and training models to maximize training effectiveness and minimize skill decay by creating training curricula customized to each surgeon's knowledge and skills. The LASTS skill model design goals are detailed enough to represent the skills needed for surgical procedures, moment-by-moment, and are sufficiently powerful to support an assessment of the skills that will differentiate experts from novices. The model accounts for variations in the skills required by different procedures by expanding on previous hierarchical models of LS skills[23] and representing the skill hierarchy of each procedure. Surgical procedures are decomposed into a tree of steps, tasks, subtasks, and where appropriate, into motions. The tree describes a choreographed sequence of motions, views, information elements, and decisions, as well as deviations for nonstandard anatomy, emergencies, or repairs for adverse events. The skill model includes psychomotor, visual–spatial, and cognitive skills, and uses annotations to represent the goals, dangers, techniques, expectations, information requirements, and situational awareness needed by the step, task, or subtask.[23] In addition to the base actions of the procedure, there may be a need to perform optional and repair actions, which are not necessary in every case. For example, a small amount of bleeding often occurs when detaching the gallbladder from the liver in LS cholecystectomy, and this may require cauterization or other repair at any point during the removal step. These action specifications also include annotations that indicate when the action should be performed.

Figure 2 shows a portion of the skill tree for the "Remove Gallbladder" step of a cholecystectomy procedure. Our skill tree graphics include the set of steps for the procedure (each page is one step, so this figure shows a single step), as well as the tasks (the black boxes) and subtasks (the white boxes) of each step. Tasks and subtasks are linked via decomposition links (solid lines) and ordering links (dotted lines with arrows). Not depicted in these graphics are the possible deviations and repair steps. Because of the diagram's space constraints and because these visualizations are meant as an overview of the procedure only, we do not show the annotations in this graphic. However, the model could be presented with an overlay of the cognitive complexity, motor complexity, or danger ratings to show how specific conditions or complications can affect the skills needed by the procedure.

SUSTAIN Critical Skills

In our ongoing research and development effort, we are supplementing the LASTS cognitive task analysis with empirical skill acquisition and retention data to inform the LASTS skill models, as well as developing and validating a SUSTAIN that includes novel, objective assessment metrics and training modules within a modular design to support training within a variety of environments, including deployed settings.

We focused initially on identifying assessment metrics that could assess LS skill acquisition, proficiency, and decay. Our initial investigations have focused on assessment of the current tasks and metrics used to train and assess LS psychomotor skills. Later investigations will leverage the outcomes of the LASTS cognitive skills identification.

The manual skills testing component of the FLS training curriculum is intended to train and measure psychomotor skill performance during basic laparoscopic surgical maneuvers; however, some visual–spatial skills are involved as well. Figure 3 provides a task description and decomposition of the constituent psychomotor and visual–spatial skills involved in each of the FLS manual skills tasks. Additionally, the most challenging aspect of each task and common strategies used were identified based on expert interviews and are included in Figure 3. Currently, the cognitive skills associated with laparoscopic surgery are trained separately within the FLS didactic program, and are assessed via a written test.

Based on skill decay literature, it is cognitive skills that decay most rapidly and therefore are most susceptible to decay during periods of nonuse, whereas psychomotor skills decay at slower rates. Therefore, it is not surprising that the FLS manual skills retention studies have shown little decay over long periods of retention. However, the current metrics for assessing the FLS psychomotor skills are limited primarily to time to complete each task, and secondarily to overt errors. Some research within the domain of robot-assisted laparoscopic surgery has begun to examine novel and objective metrics for assessing similar skills.[25,26] We suggest that more sensitive metrics may identify early indications of skill decay, and provide a means for assessing and refreshing training during and following deployments.

We propose that additional research is needed to examine the retention of the relevant perceptual and cognitive skills,

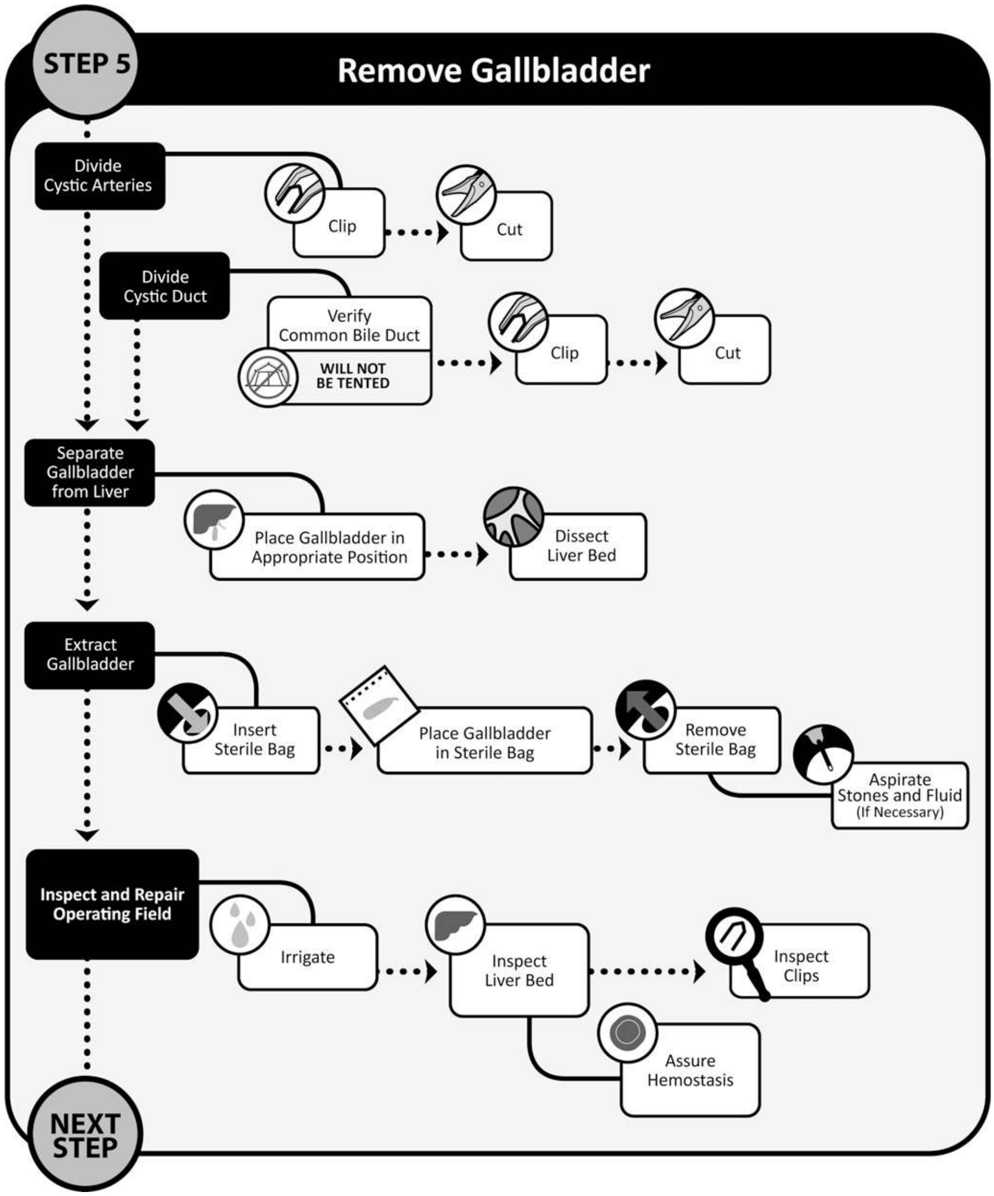


FIGURE 2. Illustration of a portion of the LASTS skill tree for a single variation of a laparoscopic cholecystectomy procedure.

and also to explore potential novel objective metrics related to psychomotor skill acquisition and decay.

SECTION 3: OBJECTIVE METRICS

To develop, validate, and effectively use models of surgical skill for training and refreshing, objective metrics are needed that can determine a surgeon's level of expertise with respect to the identified critical surgical skills. The surgical skill assessment must reliably differentiate expert surgeons from novices. An assessment that cannot make this distinction is not valid as a tool for understanding, training, and certifying surgical skills. The skill assessment metrics must provide measures of the trainee's skill level on the spectrum from novice to expert to enable the selection of training materials and practice tasks for an individual profile. Ideally, these materials and tasks will meet the surgical trainee at his or her current skill level and provide an opportunity to improve. Metrics must also effectively measure the underlying psychomotor, visual–spatial, and cognitive skills associated with LS proficiency, both independently and in conjunction with one another. Finally, the established metrics must be objective and empirically validated.

LASTS Objective Metrics

LASTS's design builds on the FLS skill measurement approach and leverages measures recorded automatically in simulation, such as speed, accuracy, or smoothness of motion based on tracking instruments and task outcome.[27,28] For example, to assess psychomotor skills, the surgeon performs

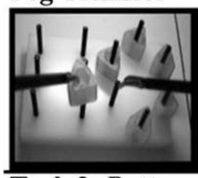
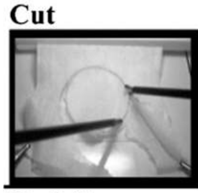
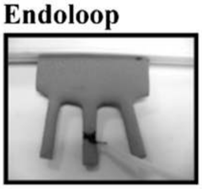
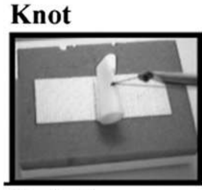
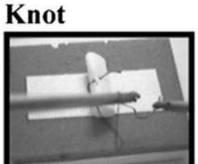
	Psychomotor Skills Involved	Visual-spatial Skills Involved	Challenging Aspects	Strategies Employed
Task 1: Peg Transfer 	Eye-hand coordination, Ambidexterity, Bimanual skills (transferring), Control precision, Manual dexterity, Arm-hand steadiness	Depth Perception, Spatial processing, Spatial relationship, Near vision, Color discrimination	Time, Ambidexterity	Strategically grasping and placing peg for security and efficiency of movement
Task 2: Pattern Cut 	Application of tension / counter-tension	Depth Perception, Near vision, Shape processing	Precision when cutting at difficult angles	Approaching from angle, Switching cutting hand or direction part-way through, Distance and angle to apply tension
Task 3: Endoloop 	Eye-hand coordination, Bimanual skills, Control precision, Manual dexterity,	Depth perception, Spatial processing, Spatial relationships	Preventing loop from tightening prior to placement on the targeted line	Tightening loop with one hand, Pressing down on targeted line with applicator while tightening, Knowing when to apply one endoloop or two
Task 4: Extracorporeal Knot 	Eye-hand coordination, Bimanual skills, Control precision, Manual and finger dexterity, Arm-hand steadiness, Approximation and not strangulation	Depth perception, Near vision, Spatial processing, Combining multiple perspectives	Proper needle grasping, Precise suture placement	Maintaining tension in both ends of thread without avulsing penrose drain, Understanding and pattern recognition of various types of knots
Task 5: Intracorporeal Knot 	Eye-hand coordination, Bimanual skills, Control precision, Manual and finger dexterity, Arm-hand steadiness, Approximation and not strangulation	Depth perception, Near vision, Spatial processing, Spatial relationships	Proper needle grasping and alignment, Precise suture placement, Use of short stitch	Beginning with suture in non-dominant hand, Creating and maintaining a short tail, Aligning heel of needle with driver the stitch is wrapping around

FIGURE 3. FLS manual skills task decomposition.

a physical task, such as suturing, using a virtual trainer. To provide a comprehensive view of surgical skills, LASTS is designed to assess not only psychomotor skills but also visual–spatial and cognitive skills (yet to be determined) as well. Visual–spatial tasks include manipulating a virtual laparoscope to obtain the proper view on an organ; finding structures, such as a gallbladder, from video of LS; and recognizing abnormal conditions, such as cirrhosis of the liver, from video of LS. Cognitive knowledge can be assessed via test questionnaires, such as the multiple choice questions used by the FLS didactic exam.[29] A number of decision tasks can be derived from the model of surgical skills for the procedure, such as where to place trocars for patients of varying body habitus (i.e., overweight, underweight, or normal weight) or whether to convert to open surgery after watching a video of a complication event. These measures of decision tasks can be formatted into multiple choice or select-on-a-diagram forms, or the measures may occur during a training procedure within the context of a simulated scenario.

In addition to assessing individual skill performance, LASTS's design also assesses performance of combined skills through tasks that incorporate two or more skills simultaneously. An example of a combined psychomotor and cognitive task is dissecting the gallbladder from the liver during a simulated LS cholecystectomy procedure. LASTS specifies that the surgeon must understand the cognitive elements of the task—the goals, risks, techniques, expectations, and situational awareness—and use those elements to modulate the psychomotor movements of the task. For example, in some cases during a specific step, the surgeon must be careful not to unintentionally puncture the gallbladder or the liver causing excessive bleeding. Finally, LASTS's design assesses combined psychomotor, visual–spatial, and cognitive skill by simulating multiple steps within a procedure. During the simulation, LASTS uses the skill tree to assess the performance of each step, task, and subtask independently. Specific scoring mechanisms are in development, which will take into account correlations between both individual and combined skills with task performance outcomes.

SUSTAIN Objective Metrics

Several objective metrics have been identified that may support effective assessment of LS skills acquisition, proficiency, and decay, including simulator-based metrics, motion- and vision-tracking metrics, and cognitive assessment metrics. For example, laparoscopic virtual environment training systems have incorporated a number of automated performance assessment metrics, such as task completion time, economy of motion, instrument collisions, peak instrument force, and peak strain. While further validation of such metrics is needed, including demonstration of predictive validity within live tissue and cadaveric models, these metrics may provide useful objective assessment and tracking tools if integrated into the FLS video trainer and/or curriculum. Laparoscopic simulation trainers can detect various motions as part of their mechanical input and visual feedback systems; however, the current FLS trainer boxes are relatively low-tech and do not include instrumentation tracking capabilities. Although the low-tech nature of the current FLS training system supports low-cost training, the goal of the SUSTAIN research and development effort is to investigate potential objective metrics that might provide insight into FLS skills acquisition, proficiency, and decay, including more sensitive metrics of psychomotor skills to enable detection of skill decay despite slower decay rates of these skills as compared to cognitive skills.

This effort seeks to develop metrics based on instrument tracking, as well as metrics that are novel within this domain, such as motion tracking via wearable instruments (i.e., gloves), and additional metrics to support assessment of perceptual and cognitive skills, both independently and in conjunction with the FLS psychomotor skills tasks.

A wide variety of motion-tracking technologies, such as instrumented gloves, currently exist for applications ranging from graphical animation to recognition of sign language. Data that may be extracted from these technologies include hand tremors, tool grip, and tool state (e.g., activating tool, rotating tool), as well as movement certainty. Instrumentation of the LS surgical tools may provide valuable data, such as time metrics, efficiency metrics, and task-specific metrics (e.g., object grasping, transfer, and placing times), as well as efficiency metrics, such as overall efficiency (left hand, right hand, average), pick-up efficiency (left hand, right hand, reverse, average), passing efficiency (left to right, right to left, average), and placing efficiency (left hand, right hand, average).

Vision tracking technologies are used in a variety of domains to provide insight into human–computer interaction, perceptual processing, and cognitive process. Vision tracking technologies have become increasingly affordable and unobtrusive, and their use has increased in both research and operational applications as a result. Vision tracking enables researchers to identify visual scan and fixation patterns during perceptual–cognitive and perceptual–motor tasks, providing insights into learning. In addition to enabling researchers to study learning patterns to develop more robust models of skill acquisition and decay, vision tracking can also enable surgical instructors and training proctors to see where trainees are focusing their visual attention during task completion in real time.

A variety of cognitive assessment technologies and strategies exist that incorporate indicators of cognitive workload and learning, such as concurrent performance of a secondary task.[11,14,15]

A subset of the metrics identified, such as hand and instrument tracking, are currently being empirically tested, and will be considered for inclusion in the SUSTAIN prototype system. In addition, we will consider cognitive assessment metrics such as assessing the ability to concurrently perform a secondary task more effectively as performance on the primary task of interest becomes more automated. While Stefanidis et al[11,14,15] demonstrated that secondary task performance provides a means for differentiating between individuals of variable LS expertise when standard performance measures fail to do so, and that training to automaticity using a secondary task correlates to improved intraoperative performance, the secondary task used in these studies was not directly related to surgical skills. It is possible that by incorporating a secondary task that addresses intraoperative skills, including cognitive skills, training would be enhanced further, providing justification for the extended time associated with training to automaticity.

SECTION 4: SKILL DECAY CURVES

Reviews of the skill decay literature have identified a core set of factors that influence the decay of trained skills. These factors include (a) length of the nonuse interval, (b) degree of overlearning (training beyond mastery), (c) task characteristics (e.g., psychomotor versus cognitive, number of steps involved), (d) method of assessing original acquisition and retention (i.e., type of test), (e) condition of retrieval (e.g., recall versus recognition), (f) instructional strategies and training methods, (g) individual differences (e.g., spatial ability), and (h) motivation.[1,2] Modeling these factors within skill decay curves helps to predict skill degradation over time and introduce refresher training at appropriate intervals.

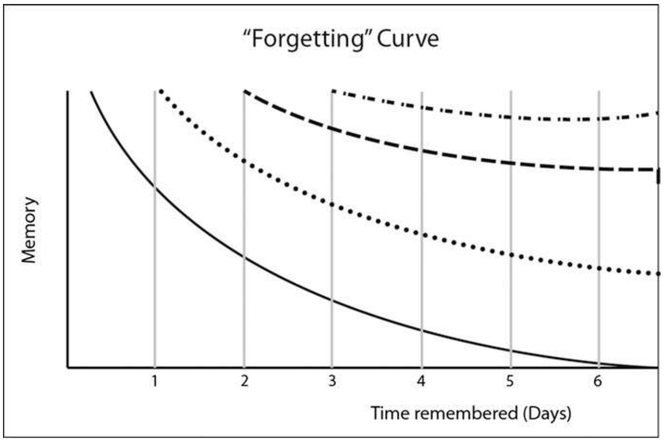


FIGURE 4. Representation of forgetting curve proposed by Ebbinghaus (1885).

Skill acquisition curves have been developed, for example, based on the relationship between the time to perform a task and the number of practice attempts.[5,6] Ebbinghaus[30] proposed the first formal decay curve, which he called the "forgetting curve." This model, shown in Figure 4, demonstrates that the rate of forgetting depends on several factors, but that typically single trial learning results in exponential skill decay, with each additional learning trial resulting in increased retention. The solid line indicates the rate of forgetting for a single learning trial; each dashed line to the right represents the rate of forgetting for additional learning trials on subsequent days.

LASTS Skill Decay Model

The skill decay models used in LASTS predict how surgical skills change over time and practice sessions, in relation to specific learning and decay factors. These models have three related parts: (1) how skills are initially acquired through training, including factors such as the order, duration, and repetition of training, as well as individual trainee differences, such as skill level and experience; (2) how skills decay over time through nonuse; and (3) how skills are reacquired through refresher courses. It is important to note that reacquisition is different from the initial acquisition because of previous exposure to the skills.[12] The model includes three skill curves for each skill assessed: (1) learning, (2) decay, and (3) relearning. The learning and relearning curves measure skill level against training attempts (i.e., number of times performing the training task). LASTS includes these factors because they have been shown by Arthur et al to modulate performance.[1] A learning curve shows how a surgeon's skill improves with each attempt.[12,31] The decay curves measure skill level over time.[11]

Learning curves for LS are typically defined in terms of the time it takes to complete a task versus how many sessions have been trained. LASTS extends the task completion time measure to all of the factors that comprise the assessment methodology to create a unique set of learning curves for each procedure, step, and task. A key contribution of LASTS is the ability to combine these curves across skills. We expect that initial learning rates (i.e., time to reach proficiency) will decrease and decay rates will increase when the surgeons are assessed on multiple skills simultaneously. Based on our initial understanding of the literature on multiple learning curves,[28] our hypothesis is that the curves will all have a similar shape because proficiency at a given skill will increase with each trial until a plateau is reached. The exact values of the curves are dependent on the skill as well as the student's experience level and past performance. On average, decay curves show decreasing proficiency over time, as the skill decays.[11]

SUSTAIN Skill Decay Model

The skill decay model proposed for SUSTAIN is both theoretical and empirical, and will provide input to the LASTS computational decay models. To begin developing an empirical model of LS skills that includes novel metrics, we conducted a pilot study[32] in which both traditional (i.e., time and error) and novel (i.e., instrumented gloved hand motion tracking) FLS assessment metrics were collected for 11 medical students with no prior exposure to the FLS training curriculum and one expert laparoscopic surgeon. FLS manual skills performance data were collected for the novice participants before training (pretest), following completion of training to proficiency (post-test), and following an 8- to 10-week retention period (follow-up). Novice spatial abilities were also measured via three validated computer tests that assessed egocentric spatial ability, allocentric spatial ability, and mental rotation. Despite a small sample size, based on traditional FLS scores, novice performance was shown to be significantly higher at post-test, as compared to pretest; FLS scores decreased at follow-up, but did not differ significantly from post-test scores. A significant difference was also found between the novice pretest and expert's scores for the five tasks, but not between the novice post-test and expert's scores, or between the novice follow-up and expert's scores; however, the novice post-test and follow-up scores would likely be significantly differentiated from expert's scores with a larger sample population. These pilot data are indicative of skill acquisition and decay trends, providing the basis for an initial empirical skill model. We also established initial technical feasibility for the use of instrumented glove motion tracking to assess smoothness of surgeon hand motions, and for the incorporation of these novel metrics within a skill decay model. Although limited by the small sample size, this pilot study provided a basis for an initial model of laparoscopic surgical skill acquisition and decay that incorporates a variety of metrics, which will be further refined and expanded to include cognitive, perceptual, and psychomotor aspects, and will be validated under ongoing empirical research efforts.

SECTION 5: DESIGN OF REFRESHER TRAINING

Stefanidis and Heniford[33] suggest that a successful laparoscopic skills curriculum should "encompass goal-oriented training, sensitive and objective performance metrics, appropriate methods of instruction and feedback, deliberate, distributed, and variable practice, an amount of overtraining, maintenance training, and a cognitive component" (p. 77).

The LASTS approach includes a curriculum generator, which will use the LASTS and SUSTAIN assessment methods, learning curves, and decay curves to create a curriculum of refresher training for a specific surgeon based on that surgeon's initial skill level across various tasks and the length of retention interval for various skills and tasks. At this time, we have designed the framework of the curriculum generator, which we describe here, but we have not yet implemented its components. The curriculum generator will leverage the skill decay curves developed, which predict the rate at which various skills will decay, and will be designed to use a surgeon's training history, history of procedures performed, and skills assessments to determine which skills are most likely to require refreshing for a specific procedure. The curriculum generator then creates a curriculum to refresh these skills. Because the curriculum focuses on refreshing the skills that have decayed for the surgeon (and not all skills), it decreases retraining time and addresses potential problems proactively. This curriculum may then be used as one part of a surgical training, certification, and recertification program. By addressing the factors of the skill decay curves that most influence decay, the curriculum generator may be able to train new surgeons and retrain experts to build robust skills that decay slowly, if at all.

For each student, the curriculum generator creates a training regimen and keeps a profile of the trainee to assess progress. The trainee can be tracked through the training program to recommend when they are ready for operating room experience, and to test this recommendation at computed intervals in collaboration with attending surgeons' review. The system first determines the trainee's skill levels using our assessment methods. It then selects training elements, such as didactics and tasks, from a library and generates new simulation training scenarios as needed. For a deploying surgeon, the system takes a last predeployment reading on skill level, and then receives the record of procedures performed during deployment. Additional information such as a surgeon's training and individual competency as rated by supervisors is collected and inputted into the curriculum generator. With this information, the curriculum generator develops a refresher course. For example, a surgeon that performed few procedures, had little access to refresher material, and had long periods of time without operations will require more extensive refresher training than a surgeon that performed many operations, refreshing their skills with related procedures.

Psychomotor and visual–spatial skills are modeled by continuous values along multiple skill dimensions, and training elements are selected to challenge the surgeon at his current estimated skill levels using item response theory. Cognitive skills are modeled as knowledge bases of correct and incorrect rules, and training elements are selected that extend these knowledge bases using intelligent tutoring systems techniques. Finally, the curriculum generator generates surgical scenarios for practice on surgical simulators. It uses narrative generation techniques[34–36] to present realistic scenarios that also accomplish training goals. LASTS determines which skills need refreshing by using the decay curves and continuous skill assessments.

To augment the individual training elements, LASTS uses a scenario-based training methodology, where surgical situations are presented to the trainees instead of abstract tasks. This approach is designed to enable a higher level of skill acquisition and decrease decay. Techniques in search-based interactive narrative and narrative planning models are particularly suited for this type of generation, since they can be used to generate a sequence of realistic scenarios (the narrative content) while optimizing the ability for the student to learn the content. Narrative planning approaches provide this functionality.[32] The role of the narrative planning model is to create realistic, scenario-based training sessions that cover the skills in the necessary quantity and order.

The LASTS computational models and curriculum generator will be incorporated into the SUSTAIN prototype, which includes a modular design to support training within a variety of environments, including deployed settings.

SUMMARY AND CONCLUSIONS

Although the research literature on skill acquisition and decay is comprehensive and extensive, this is not true of the literature that pertains specifically to LS skill decay. Based on the current research literature, it is evident that a wide variety of factors related to task and training characteristics, retention intervals, and transfer task environments impact the rate of decay of various skill sets, and that subcomponents of skill sets decay at different rates. However, relatively little is known about the nature of LS skill retention. Empirical research is needed to develop and validate predictive skill decay models related to these highly specialized skills. This research will help provide guidelines for efficient and effective refresher training that can be implemented within standardized medical training curricula, particularly for military surgeons, to prevent laparoscopic skill deterioration during long deployment cycles.

The research questions yet to be addressed in this area are many: Are there different decay rates for different types of learning? Which instructional methods/strategies are the most effective for skill/knowledge acquisition to prevent skill decay? What are the most effective strategies for refresher training (reacquisition)? How do individual factors, such as spatial ability and handedness, impact skill decay?

Studies assessing decay of LS skills often do not assess cognitive skills or knowledge, which predicts decay. Also, the studies that address skill acquisition and decay suffer from a lack of objective, reliable, and valid measures of skill level and decay, as well as a lack of data on reacquisition of LS skills. When acquisition is studied, assessments are usually in relation to new training techniques compared to a traditional method, and it is difficult to interpret and draw conclusions from these studies without objective, reliable, and valid metrics. Further, many of the studies assessing acquisition and retention are based on psychomotor tasks, and do not address cognitive or perceptual skills. To remedy this situation, we suggest the following approach: (1) develop and identify assessment metrics that are objective, reliable, and valid to assess these skills over time; (2) acquire a comprehensive understanding of the nature of laparoscopic surgical skill acquisition and decay, including identification of the requisite critical skills; (3) report skill decay curves that enable prediction of which critical skills decay and at what rate; (4) develop training strategies for simulation-based training that support rapid acquisition and long-term retention; and (5) develop retraining for sustainment of these critical and perishable skills.

An ongoing research and development effort has been funded by the Office of Naval Research and the Telemedicine and Advanced Technology Research Center to further the development and empirical validation of novel objective metrics for laparoscopic surgical skill acquisition, decay, and reacquisition model; design, develop, and validate a prototype simulation-based assessment and training module; and develop curriculum recommendations for the prevention of LS skills decay. The goal of this effort is to conduct research on LS to develop generalized methods, procedures, and training methods that could be used to prevent decay of other critical medical skills and knowledge.

ACKNOWLEDGMENTS

This work is being conducted in collaboration with the Uniformed Services University of the Health Sciences, the National Capital Simulation Center, and Harvard Medical School. This research is funded by the Office of Naval Research and the Telemedicine and Advanced Technology Research Center under two Small Business Innovation Research contracts, AnthroTronix, Inc. (N0001411C0420) and Charles River Analytics, Inc. (N0001411C0426). The work reported herein was also partially supported by a grant from the Office of Naval Research, Award Number N00014-10-1-0978.

REFERENCES

1. Arthur W, Bennett W, Stanush PL, McNelly TL: Factors that influence skill decay and retention: a quantitative review and analysis. Human Performance 1998; 11(1): 57–101.
2. Wisher RA, Sabol MA, Ellis J: Staying sharp: retention of military knowledge and skills (Rep. No. ARI Special Report 39). Alexandria, VA, U.S. Army Research Institute for the Behavioral and Social Sciences, 1999. Available at http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA366825; accessed May 7, 2013.
3. Brunner WC, Korndorffer JR: Laparoscopic virtual reality training: are 30 repetitions enough? J Surg Res 2004; 122(2): 150–6.
4. Jacklin R, Sevdalis N, Darzi A, Vincent CA: Efficacy of cognitive feedback in improving operative risk estimation. Am J Surg 2009; 197(1): 76–81.
5. Newell A, Rosenbloom PS: Mechanisms of skill acquisition and the law of practice. In: Cognitive Skills and Their Acquisition, pp 1–55. Edited by Anderson JR, Hillsdale, NJ, Erlbaum, 1981.
6. Heathcote A, Brown S, Mewhort DJK: The power law repealed: the case for an exponential law of practice. Psychon Bull Rev 2000; 7(2): 185–207.
7. Fitts P, Posner M: Human Performance. Belmont, CA, Brooks/Cole, 1967.
8. Fried GM, Feldman LS, Vassiliou MC, et al: Proving the value of simulation in laparoscopic surgery. Ann Surg 2004; 240(3): 518.
9. McCluney AL, Vassiliou MC, Kaneva PA, et al: FLS simulator performance predicts intraoperative laparoscopic skill. Surg Endosc 2007; 21(11): 1991–5.
10. Soper NJ, Fried GM: The fundamentals of laparoscopic surgery: its time has come. Bull Am Coll Surg 2008; 93(9): 30.
11. Stefanidis D, Scerbo MW, Montero PN, Acker CE, Smith WD: Simulator training to automaticity leads to improved skill transfer compared with traditional proficiency-based training: a randomized controlled trial. Ann Surg 2012; 255(1): 30.
12. Stefanidis D, Korndorffer JR, Markley S, Sierra R, Scott DJ: Proficiency maintenance: impact of ongoing simulator training on laparoscopic skill retention. J Am Coll Surg 2006; 202(4): 599–603.
13. Hiemstra E, Kolkman W, van de Put MAJ, Jansen FW: Retention of basic laparoscopic skills after a structured training program. Gynecol Surg 2009; 6(3): 229–35.
14. Stefanidis D, Scerbo MW, Korndorffer JR, Scott DJ: Redefining simulator proficiency using automaticity theory. Am J Surg 2007; 193(4): 502–6.
15. Stefanidis D, Scerbo MW, Sechrist C, Mostafavi A, Heniford BT: Do novices display automaticity during simulator training? Am J Surg 2008; 195(2): 210–3.
16. Park SH, Suh IH, Chien JH, Paik JH, Ritter FE, Siu KC: Modeling surgical skill learning with cognitive simulation. In: Medicine Meets Virtual Reality 18. Edited by Westwood JD, Westwood SW, Felländer-Tsai L, et al. Amsterdam, The Netherlands, IOS Press, 2011.
17. Palter VN: Comprehensive training curricula for minimally invasive surgery. J Grad Med Educ 2011; 3(3): 293–8.
18. Hassan I, Gerdes B, Koller M, et al: Spatial perception predicts laparoscopic skills on virtual reality laparoscopy simulator. Childs Nerv Syst 2007; 23(6): 685–9.
19. Ritter EM, McClusky DA III, Gallagher AG, Enochsson L, Smith CD: Perceptual, visuospatial, and psychomotor abilities correlate with duration of training required on a virtual-reality flexible endoscopy simulator. Am J Surg 2006; 192(3): 379–84.
20. Gallagher AG, Cowie R, Crothers I, Jordan-Black JA, Satava RM: PicSOr: an objective test of perceptual skill that predicts laparoscopic technical skill in three initial studies of laparoscopopic performance. Surg Endosc 2003; 17(9): 1468–71.
21. Keehner M, Lippa Y, Montello DR, Tendick F, Hegarty M: Learning a spatial skill for surgery: how the contributions of abilities change with practice. Appl Cogn Psychol 2006; 20(4): 487–503.
22. Keehner MM, Tendick F, Meng MV, et al: Spatial ability, experience, and skill in laparoscopic surgery. Am J Surg 2004; 188(1): 71–5.
23. Cao CGL, MacKenzie CL, Ibbotson JA, et al: Hierarchical decomposition of laparoscopic procedures. In: Medicine Meets Virtual Reality: The Convergence of Physical and Informational Technologies: Options for a New Era in Healthcare. Edited by Westwood JD, Hoffman HM, Robb RA, Stredney D., Amsterdam, The Netherlands, IOS Press, 1999.
24. Scott-Connor CEH: The SAGES Manual, Fundamentals of Laparoscopy, Thoracoscopy, and GI Endoscopy. New York, NY, Springer Science + Business Media, 2006.
25. Chang L, Satava RM, Pellegrini CA, Sinanan MN: Robotic surgery: identifying the learning curve through objective measurement of skill. Surg Endosc 2003; 17(11): 1744–8.
26. Verner L, Oleynikov D, Holtmann S, Haider H, Zhukov L: Measurements of the level of surgical expertise using flight path analysis from Da Vinci robotic surgical system. Stud Health Technol Inform 2003; 94: 373–8.
27. O'Connor A, Schwaitzberg SD, Cao CGL: How much feedback is necessary for learning to suture? Surg Endosc 2007; 22: 1614–9. Available at

http://link.springer.com/content/pdf/10.1007%2Fs00464-007-9645-6.pdf; accessed May 7, 2013.

28. Zhou M, Tse S, Derevianko A, Jones DB, Schwaitzberg SD, Cao CGL: The effect of haptic feedback on laparoscopic suturing and knot-tying: a learning curve study. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2008. Santa Monica, CA, Human Factors and Ergonomics Society, 2008.
29. Peters J, Fried GM, Swanstrom LL, et al: Development and validation of a comprehensive program of education and assessment of the basic fundamentals of laparoscopic surgery. Surgery 2004; 135(1): 21–7.
30. Ebbinghaus H: Über das gedächtnis: Untersuchungen zur experimentellen psychologie. Leipzig, Germany, Duncker and Humboldt, 1885. [English translation: Memory: A Contribution to Experimental Psychology. New York, NY, Dover Publications, 1964.]
31. Korndorffer JR, Dunne JB, Sierra R, Stefanidis D, Touchard C, Scott D: Simulator training for laparoscopic suturing using performance goals translates to the operating room. J Am Coll Surg 2005; 201(1): 23–9.
32. Skinner A, Lathan C: Assessment of laparoscopic surgical skill acquisition and retention. In: Studies in Health Technology and Informatics: Medicine Meets Virtual Reality 19. Edited by Westwood JD, Westwood SW, Fellander-Tsai L, et al. Amsterdam, The Netherlands, IOS Press, 2012.
33. Stefanidis D, Heniford BT: The formula for a successful laparoscopic skills curriculum. Arch Surg 2009; 144(1): 77.
34. Niehaus J, Riedl MO: Scenario adaptation: an approach to customizing computer-based training games and simulations. In: Proceedings of the AIED 2009 Workshop on Intelligent Educational Games 2009. Brighton, UK, AIED, 2009. Available at http://www.cc.gatech.edu/~riedl/pubs/aied-ieg09.pdf; accessed May 7, 2013.
35. Riedl M, Young RM: An intent-driven planner for multi-agent story generation. Proceedings of the Third International Conference on Autonomous Agents and Multi Agent Systems, July 2004.
36. Weyhrauch P: Guiding interactive drama. School of Computer Science, Carnegie Mellon University, Doctoral Thesis, Pittsburgh, PA, 1997.