Adaptive and Perceptual Learning Technologies in Medical Education and Training

Philip J. Kellman, PhD

ABSTRACT Recent advances in the learning sciences offer remarkable potential to improve medical education and maximize the benefits of emerging medical technologies. This article describes 2 major innovation areas in the learning sciences that apply to simulation and other aspects of medical learning: Perceptual learning (PL) and adaptive learning technologies. PL technology offers, for the first time, systematic, computer-based methods for teaching pattern recognition, structural intuition, transfer, and fluency. Synergistic with PL are new adaptive learning technologies that optimize learning for each individual, embed objective assessment, and implement mastery criteria. The author describes the Adaptive Response-Time-based Sequencing (ARTS) system, which uses each learner's accuracy and speed in interactive learning to guide spacing, sequencing, and mastery. In recent efforts, these new technologies have been applied in medical learning contexts, including adaptive learning modules for initial medical diagnosis and perceptual/adaptive learning modules (PALMs) in dermatology, histology, and radiology. Results of all these efforts indicate the remarkable potential of perceptual and adaptive learning technologies, individually and in combination, to improve learning in a variety of medical domains.

INTRODUCTION

Recent advances in the learning sciences offer remarkable potential to improve medical education. These advances are relevant to almost all domains of medicine, and they have direct application to maximizing the benefits of simulation and cutting-edge technologies. In this article, I describe two innovations in training technology that apply to simulation and other aspects of medical learning: perceptual learning (PL) and adaptive learning technologies. PL techniques teach pattern recognition, structural intuition, and fluency. Adaptive learning technologies can optimize learning for each individual, embed objective assessment throughout learning, and implement mastery criteria.

Understanding the role and value of these emerging technologies requires some discussion of traditional conceptions of learning and how these are changing, as well as elaboration of basic elements and benefits of each technology. We consider conceptions of learning, perceptual learning technology, and adaptive learning technology in the first three sections. Then, we describe recently developed medical learning applications of perceptual and adaptive learning technology, in the areas of clinical diagnosis, radiology, dermatology, and histopathology. In the final section, we consider synergies between these learning technologies and simulation tools and techniques in medicine.

Department of Psychology, University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, CA 90095-1563.

Any opinions, findings, and conclusions or recommendations expressed in this article are those of the author and do not necessarily reflect the views of the U.S. Department of Education, the National Science Foundation, or other agencies. The findings and opinions expressed here do not necessarily reflect the positions or policies of the Office of Naval Research. Systems that use learner speed and accuracy to sequence learning events, as well as some aspects of perceptual learning technology described here, are protected by U.S. patent 7052277 and patents pending, assigned to Insight Learning Technology, Inc. For information, please contact either the author or Info@insightlearningtech.com.

doi: 10.7205/MILMED-D-13-00218

REVISITING LEARNING

In most instructional settings, learning is organized around two types of knowledge. This is not surprising, as these two types are often considered exhaustive, even in many cognitive psychology texts. *Declarative knowledge* includes facts and concepts that can be verbalized. *Procedural knowledge* includes sequences of steps that can be enacted. A conventional view of learning, shared by nonspecialists and researchers alike, is that learning consists of accumulating these facts, concepts, and procedures.[1] The standard view has been called a "container" model of the mind: Learning consists of facts, concepts, and procedures that we place into the container (the mind), and for later performance, we retrieve these items.[1]

Persistent problems in learning and instruction suggest that this learning worldview is defective. Students who have been faithfully taught and have diligently absorbed declarative and procedural inputs fail to recognize key structures and patterns in real-world tasks. Students may know procedures but fail to understand their conditions of application or which ones apply to new problems or situations. And, learners may understand but process slowly, with high cognitive load, making them impaired in demanding, complex, or time-limited tasks.

These characteristic problems can be observed in learning domains from mathematics to surgical training. They suggest that much is missing from the typical view of learning. What is it? Some answers are clearly available if one looks, not at the literatures on education or learning, but the literature on *expertise*. Studies of expertise—what people are like when they are really good at things—recurrently implicate a number of abilities that emerge from changes in the

TABLE I. Some Characteristics of Expert and Novice Information Extraction

	Novice	Expert
Discovery Effects[a]		
Selectivity	Attention to Relevant and Irrelevant Information	Selective Pickup of Relevant Information/Filtering
Units	Simple Features	"Chunks"/Higher-Order Relations
Fluency Effects[b]		
Search Type	Serial Processing	More Parallel Processing
Cognitive Load	High	Low
Speed	Slow	Fast

[a]Discovery effects involve learning and selectively extracting features or relations that are relevant to a task or classification. [b]Fluency effects involve coming to extract relevant information faster and with lower attentional or cognitive load. (See text.)

way information is extracted: PL. Kellman[2] suggested that PL effects fall into two broad categories, *discovery* and *fluency* effects. Table I summarizes a number of these in each category. Discovery effects refer to learners finding the information that is most relevant to a task. One important discovery effect is increased attentional selectivity. With practice on a given task, learners come to pick up the relevant information for relevant classifications while ignoring irrelevant variation.[3] Practice also leads learners to discover invariant or characteristic relations that are not initially evident (cf., Chase and Simon[4]) and to form and process higher level units (Goldstone[5]; for reviews, see Bereiter and Scardamalia,[1] Gibson,[3] and Goldstone[6]).

Fluency effects refer to changes in the efficiency of information extraction. PL leads to fluent and sometimes automatic processing,[7] with automaticity in PL defined as the ability to pick up information with little or no sensitivity to task load. As a consequence, perceptual expertise may lead to more parallel processing and faster pickup of information.

It is fair to say that studies of expertise have done more to describe these characteristics of experts than to reveal how these changes come about, except for the observation that expertise grows over long experience.[8] More foundational work suggesting how these changes arise was done by Eleanor Gibson[3] and her students several decades ago. Gibson defined PL as "changes in the pick up of information as a result of practice or experience" and argued that such changes tended to be domain-specific improvements, resulting from classification experience, involving the discovery of characteristic or invariant properties distinguishing objects or situations from one another.[3]

Recently, PL has become a major focus of research in cognitive science and neuroscience (for reviews, see Kellman,[2] Fahle and Poggio,[9] and Kellman and Garrigan[10]). For present purposes, 3 clear ideas are most relevant. First, PL is a pervasive process of learning that serves to optimize information extraction to improve task performance. Second, with appropriate procedures, all kinds of feature and pattern extraction can be improved by using PL. Third, these improvements are often dramatic, sometimes improving task performance by orders of magnitude.

One example of a complex task in which dramatic PL effects have been studied is chess. On a good day, the best human chess grandmaster can beat a chess-playing computer that examines upward of 200 million possible moves per second and incorporates methods for evaluating positions and strategies culled from grandmaster consultants. By comparison, human players do relatively little raw search in chess, examining perhaps as many as 4 possible moves and following these to a depth of several successive possible moves. Despite this huge discrepancy in search ability, humans can play chess at astonishingly high levels. Remarkably, the incredible abilities of skilled chess players, relative to novice players, turn out not to depend primarily on sophisticated reasoning or a greater storehouse of factual knowledge. They depend on perception of structure: learned pattern classification abilities of remarkable flexibility, complexity, and sophistication.[4,11] Much of the relevant perception of structure is not verbally accessible. With appropriate learning experiences in a specific domain, PL allows humans to reach almost magical levels of expertise, but the relevant learning experiences are not those of traditional classrooms or tutorials. These observations about the origins of advanced expertise apply to many high-level domains of human competence; in medicine, they are crucial for understanding the skills of the expert radiologist, pathologist, and surgeon.

Likewise, PL appears to form the core of the notion of "situation awareness," which can be described as "being aware of what is happening around you to understand how information, events, and your own actions will affect your goals and objectives."[12] Situation awareness is crucially important to many domains of military training and performance, as well as aviation and air traffic control, and many other complex tasks. The PL effects given in Table I summarize much of what is involved: selectively and automatically picking up task-relevant information, detecting important relationships, and being able to extract information with low-enough cognitive load to allow handling of complex and overlapping task demands.

In these and other domains, there is a common misconception about PL effects in expertise, related both to the oft-repeated maxim that becoming an expert requires 10,000 hours of practice and the typical view of learning as storing something in the mind. The misconception is that what happens in the transition from novice to expert has to do with committing to memory a great number of examples. A related idea is the suggestion that stored instances somehow become "mental models." In chess, for example, it may be asserted that the experts succeed because they have memorized many games. These ideas do not provide a workable account of the expertise furnished by PL. Although experiencing many instances can be an important input to PL, storage of instances does not produce much of the relevant expertise, nor is it a component of leading computational models of PL (see Kellman and Garrigan,[10] for a recent review).

The reason involves what is needed to effectively use any facts, procedures, or models stored in memory (especially if there is a lot stored). Effective performance relies crucially on *pattern recognition*. When faced with a new situation, the question is: Which of the items, procedures, or models stored in the brain is relevant to *this* situation? This is a problem that requires classifying the new input. PL is the learning process that ultimately, through changes in the attunement, scope, and fluency of information extraction,[3,10] distinguishes the expert from the novice who does not see what is relevant or who is blind to the distinguishing features that place the input into one category rather than another.

In domains that matter, this can hardly ever be done by use of memorized instances. The skilled radiologist, for example, must detect the pathology in a *new* image or set of images, where the tumor may be manifest in a different location, size, orientation, and contrast, and situated amidst novel and variable background anatomy and image noise, as compared with any images seen previously. The power of exposure, classification, and feedback involving a wide variety of cases is that information selection and pattern discovery mechanisms are honed, allowing the pickup of relevant structures, and equally important, that the information extraction mechanisms discard or ignore irrelevancies that do not drive important classifications. The expert emerges from PL experience with an attuned information extraction system, not a storehouse of memorized instances. The access to relevant stored information can work effectively in complex domains only after the input is rapidly and accurately classified.

It is paradoxical (or instructive) that one encounters the instance memorization account in reference to chess, as this is a domain in which the futility of memorizing can be shown by quantitative proof, based on the fact that the sheer number of possibilities dwarfs any capacity to remember and replay specific games. It has been calculated that after 40 moves of a game of chess, there are about 10^{120} different possible games. This exceeds by a considerable amount the number of atoms in the universe (about 10^{80})! Even chess-playing machines, whose memory capacity far exceeds humans, both in volume and accuracy, are not able to play chess primarily by looking up familiar games.

PERCEPTUAL LEARNING TECHNOLOGY

Most recent PL research has focused on low–level sensory discriminations.[9] This focus derives from an interest on understanding plasticity in the brain, and from the fact that sensory coding is best understood in the early cortical levels of the brain. Considerable research, however, indicates that PL is equally applicable to high-level, complex tasks.[3,4,13–15] Many of these research efforts in both high- and low-level PL have led to an improved understanding of the conditions that produce PL.

These developments are significant, because conventional instructional techniques do little to advance expert pattern recognition and fluency. In many domains, there has been a tacit assumption that we cannot teach this kind of knowing. In accord with this assumption, radiologists, surgeons, and pathologists, as well as chemists, pilots, and air traffic controllers, are told that expert intuitions will arise, not from "book learning," but from "seasoning," "experience," or the passage of time.

From the standpoint of cognitive science, the passage of time is not a strong candidate for a learning mechanism. Instead it turns out that this kind of learning can be systematically addressed and accelerated using appropriate computer-based instantiations of principles of PL.[13,15] We call these *perceptual learning modules* (PLMs).

A complete description of PL techniques is beyond the scope of this article, but a few basics will serve to characterize the approach. PLMs use interactive learning trials; learning advances through many short trials in which the learner performs some classification task and receives feedback. Classification episodes are the engine that drives PL processes to discover and process fluently key features and relationships relevant to the task. Equally crucial are specific kinds of variation in the display sets. Instances never or seldom repeat. Positive instances must vary in characteristics irrelevant to the classification, to allow learning of invariances. Negative instances must share with positive instances the values and dimensions of irrelevant properties. Research suggests a number of other important considerations about trial formats, spacing, and sequencing.[16] The key to understanding PLMs, relative to traditional instructional modes, is that in PLMs one is seldom asked to solve an explicit problem or give a declarative answer; rather the tasks in PLMs call upon the learner to classify, locate, distinguish, or map structure across multiple representations.

Work with PLMs shows that relatively brief interventions can produce large learning gains in many domains. Some examples include aviation training,[13] mathematics,[14] and science learning.[17] In some especially novel applications, PLMs are being used to improve intuitions about patterns that may lead to drug discovery in the pharmaceutical industry.

A number of studies indicate the role of perceptual structure in science, technology, engineering, and mathematics (STEM) learning domains,[18] as well as the potential of PL interventions to accelerate expert information extraction and fluency in mathematics.[10,14,16,17,19,20] PL interventions seem to be able to overcome pervasive obstacles in mathematics learning. In a recent series of PLMs targeting interrelated concepts in linear and area measurement, units, fractions, multiplication, and division, middle school students using PLMs in targeted interventions consistently showed strong and long-lasting learning gains on assessments including primarily transfer items, with effect sizes in the range from 0.84 to 2.69.[14,16] PLM techniques systematically address aspects of expertise for which direct instructional methods have not been previously available.

Both in terms of the breadth of applications and the possibility of radically improving learning in high-stakes domains, no area is more promising for PL technology than medical learning. Within radiology alone, there are a huge number of perceptual classifications relating to classification of pathology and normal variation, spanning not only a variety of disease conditions but also several different imaging modalities. Some of these involve a small number of fixed views; others involve 3D models capable of generating many views and requiring perceptual exploration to process fully, whereas in still others, such as ultrasound, the crucial information is often available only in animated sequences. It is well known that the speed and accuracy of the expert radiologist in exploring, seeing, and classifying develop over long and unsystematic experience (and may be highly variable across individuals). Likewise, pathologists must distinguish and classify different tissue conditions and pathogens, and dermatologists must classify skin conditions. Nor is PL confined to visual displays; heart sounds and breathing abnormalities are auditory examples, and we could enumerate haptic and tactile examples as well.

Equally important are the perceptual–procedural combinations required in surgical and interventional procedures. Although we are accustomed to thinking about the deft hands of the surgeon, the crucial role of perceptual expertise in guiding procedures, recognizing tissues and organs, and providing feedback from action illustrates Benjamin Franklin's astute observation: "The eye of the master will do more work than both of his hands."

We have begun to engineer PL technology into a number of domains of medical learning, and the potential appears limitless. Before describing these initial efforts, it will be useful to introduce the companion innovation that allows us to get the most from these efforts in PL interventions and also improves the efficiency of other types of learning: adaptive learning technology.

ADAPTIVE LEARNING TECHNOLOGY

In most instructional settings, student learning is limited by the failure of instruction to adapt to the individual. Students have different starting points and differ in aspects of lessons they learn well or poorly. Testing often arrives at the end, not in the midst, of learning, and it often involves global scoring rather than rich descriptions of what has and has not been learned. Moreover, testing usually targets accuracy alone, or perhaps speed for an entire test. Seldom are combined accuracy and fluency measures used to assess detailed aspects of learning; nor are assessments fed back continuously to optimize each individual's learning. Lacking such links between continuous assessment and the flow of learning events, it is also rare for the learner to be guided to mastery criteria involving accuracy and fluency for all components of learning tasks. These limitations can potentially be overcome, and learning dramatically improved, by the use of adaptive learning technology.

Adaptive Response-Time-Based Sequencing (ARTS) System

Since the classic work of Atkinson in the 1960s,[21] a variety of adaptive learning schemes have been proposed, with the goal of using the learner's performance along with laws of learning and memory to make learning more efficient. These systems have usually been tested with the learning of discrete items, such as foreign language vocabulary words, and have been shown to outperform random presentation of items. Most systems adapt the presentation of items based on the learner's accuracy on previous trials, and some guide learning by algorithms that derive estimates of probabilities of items becoming well-learned, based on models of learning.[22,23]

The success of previous adaptive learning systems suggests the overall promise of adaptive approaches. Existing systems, however, have important limitations. One is that model-based systems require a prior experiment, using similar learners and random presentation of learning materials, to estimate parameters for implementing the adaptive scheme. Another is that reliance on accuracy omits important information that may be provided by response times (RTs).

We have developed a new adaptive learning system that uses both accuracy and speed to determine the spacing and sequencing in learning, as well as in implementing mastery criteria. We call it ARTS – Adaptive Response-Time-Based Sequencing.[24] We describe some basics of the system and then describe its utility.

Consider a set of n items (facts, patterns, concepts, procedures) to be learned. How can we optimize learning of the set for the individual learner? We assume an interactive learning system, in which learning consists primarily of learning trials. On each trial, some item, problem, or situation is presented, and the user must process and make a response. We optimize learning by applying principles of learning to a number of items simultaneously in a *priority score* system, in which all items (or categories in category sequencing) are assigned scores indicating the relative importance of that item appearing on the next learning trial. Priority scores for each item are updated after every trial, as a function of learner accuracy and RTs, trials elapsed, and in view of mastery criteria. Learning strength is assessed continuously and in some implementations, cumulatively, from performance data. In most applications, the sequencing algorithm chooses the highest priority item on each learning trial. Adjustable parameters allow flexible and concurrent implementation of principles of learning and memory, such as stretching the retention interval automatically for each item as learning strength grows.

Our system relies on a database that stores all categories in PL and all instances in factual learning contexts (e.g., multiplication facts, vocabulary, chemical symbols, etc.). Performance data for every trial and every category or instance are acquired and used by a sequencing algorithm. For simplicity, we describe the system in terms of item sequencing, although it applies also to category learning, in which each presentation involves a novel instance. Another simplification is that even in basic factual learning, multiple formats may be used across trials to test a single item (to produce generalizable learning and enhance interest), but we omit further details.

We describe aspects of the system here omitting mathematical and technical detail. (See Mettler et al[24] for more information.) Our framework has great flexibility and may use a variety of equations relating elapsed time or trials, accuracy, and RT to the priority for presentation. When any particular function of these variables is used, there are parameters that may be adjusted to suit particular learning contexts or even individual learners. Priority scores for items are dynamically updated after each trial. In many applications, initial priority scores are given to all items, and an item's score does not change until after it is first selected for presentation. This establishes a baseline priority for feeding in new items that may be balanced against changing priorities for items already introduced. Preset orderings in learning can be accomplished by the assignment of initial priority scores that are higher for some items or categories than for others.

The full set of learning principles and objectives that may be embedded in ARTS is too extensive to describe here, but some important ones include:

Rapid Item or Category Reappearance After Errors

Errors result in assignment of a high-priority weighting. With ordinary settings, the error weighting will exceed all initial priority score assignments, as well as the highest priority that may result from a slow, correct answer. However, reappearance of missed items is still subject to enforced delay.

Interleaving/Enforced Delay

To prevent recurrence of an item while its answer remains in working memory, the system is normally configured to preclude the presentation of the same item on consecutive trials.

Joint Optimization for the Entire Learning Set

A priority score system allows joint satisfaction of a number of learning principles applied to an entire set of items, as all factors feed into a priority score for each item or category. Scores are dynamically updated after each trial, and items or categories compete for selection on each learning trial.

Retirement and Mastery Criteria

Adaptive learning focuses the learner's effort where it is needed most. Commonly, learning effort and time are limited; therefore, it often makes sense to prioritize. We use the term *retirement* to describe removal of a learning item or category from the learning set, based on attainment of mastery criteria. Pyc and Rawson[25] used the term "dropout" for this idea and found evidence that greater learning efficiency can be achieved with this feature, especially in highly demanding learning situations. RTs provide important clues to the type of processing the learner is using. When a learner answers a problem by calculating or reasoning, they will tend to be slower than when retrieving the answer from memory. A key effect of PL, for example, is becoming able to extract relevant structure with low attentional load, which is an important contributor to expertise in many domains. These are independent reasons for using RT in mastery criteria.

Dynamic Spacing Based on RTs

In our system, the priority for re-presentation of an item is a function of RT and accuracy. Even with an accurate answer, a long RT suggests relatively weak learning strength. The system can use various functions of RT but typically produces increasing priority for longer RTs. Use of RTs in adaptive learning offers a simple, direct framework for implementing important principles to produce efficiencies in learning. We hypothesize an internal variable of learning strength that may be influenced by the arrangement of learning events and inferred to some degree from performance. Learning strength is reflected in accuracy and speed in generating a factual answer or in making a classification in PL. Evidence supports response speed as an indicator of learning strength.[25,26] Considerable research suggests that the value of a test trial (with successful retrieval) varies with an item's learning strength.[27,28] Thus, the best time to re-present an item is at the longest interval for which a correct retrieval can still be accomplished.[29]

Controversy persists about whether and when expanding the retention interval is superior to schedules with equal spacing.[27,30,31] Although these issues are subjects of continuing research, considerable evidence supports the idea that difficulty of successful retrieval is an important factor.[27,28,32] Pyc and Rawson[28] labeled this idea the "retrieval effort hypothesis": more difficult, but successful, retrievals are more beneficial to learning. In recent work, they studied the relation of number of successful retrievals to later memory performance, while manipulating the difficulty of those retrievals in terms of number of intervening trials. Greater numbers of intervening trials predicted better retention. These investigators also provided evidence that, as had been suggested in other work, larger gaps produced longer average response latencies,[28] a finding consistent both with the idea that a larger gap affects an item's learning strength and that learning strength is reflected in RTs. Other recent research provides evidence for a substantial advantage of expanding the retrieval interval when material is highly susceptible to forgetting or when intervening material is processed between testing events,[29] conditions that apply to many formal learning situations, including most medical learning applications. The flexibility of parameter adjustment in the ARTS system makes it possible to accommodate varied conditions of learning and even new findings regarding optimal spacing relations.

Multipurpose, Multilevel Assessment

ARTS offers not only new opportunities to improve learning but also wide-ranging possibilities for assessment. At the core of adaptive learning is performance tracking and adjustment based on embedded assessment. In our system, every concept and item in the database is tracked in terms of the learner's accuracy and RTs on past trials. Both the raw data and derived measures are continuously available to gauge a learner's progress. Aggregating across learners can show a class's strengths and weaknesses for different categories of learning.

Recent research shows that ARTS outperforms random presentation[33] and also outperforms a classic adaptive learning system[22] in tasks involving learning of factual items.[24] Other research indicates that ARTS improves learning in perceptual and category learning relative to other schemes.[33]

MEDICAL APPLICATIONS OF PERCEPTUAL AND ADAPTIVE LEARNING TECHNOLOGIES

We have begun applying perceptual and adaptive learning technologies to medical learning, and the results are remarkably promising. We describe four of these efforts briefly.

ARTS Technology for Optimal Sick Call Performance

In a recent project funded by U.S. Army RDECOM, (a collaboration of UCLA, Insight Learning Technology, and Pelagique, Inc.), we used ARTS in prototype learning systems for learning of factual material and medical diagnosis. The focus was on initial clinical "sick call" diagnosis by corpsmen and medics, and the goal was to improve factual learning through adaptive factual learning modules and integration of probabilistic information in diagnosis in cognitive task modules.

In an efficacy study using premedical students carried out in the UCLA Human Perception Laboratory, the ARTS-based factual learning modules produced highly effective learning of medical material (such as signs and symptoms of meningitis, supraglottitis, etc.) and outperformed a control group using conventional study methods.[34] Moreover, the cognitive task modules, which aimed at training information integration and higher level pattern recognition in diagnosis, added substantial benefits beyond mastery of the basic factual information.

PL in Radiology

Radiological diagnosis includes many domains in which subtle perceptual discriminations must be made, and radiological training could likely be radically improved by appropriate deployment of perceptual and adaptive learning technology. In a pilot project, we have begun to apply these methods to X-ray diagnosis of wrist injuries. Figure 1 shows a sample screenshot. A variety of trial types, including distinguishing normal from injured wrists and classification of single or multiple injuries in particular images, are used in the module to maximize PL. Studies are ongoing, but initial results suggest that this format for learning can produce strong advances in perceptual expertise from relatively short investments of learning time.

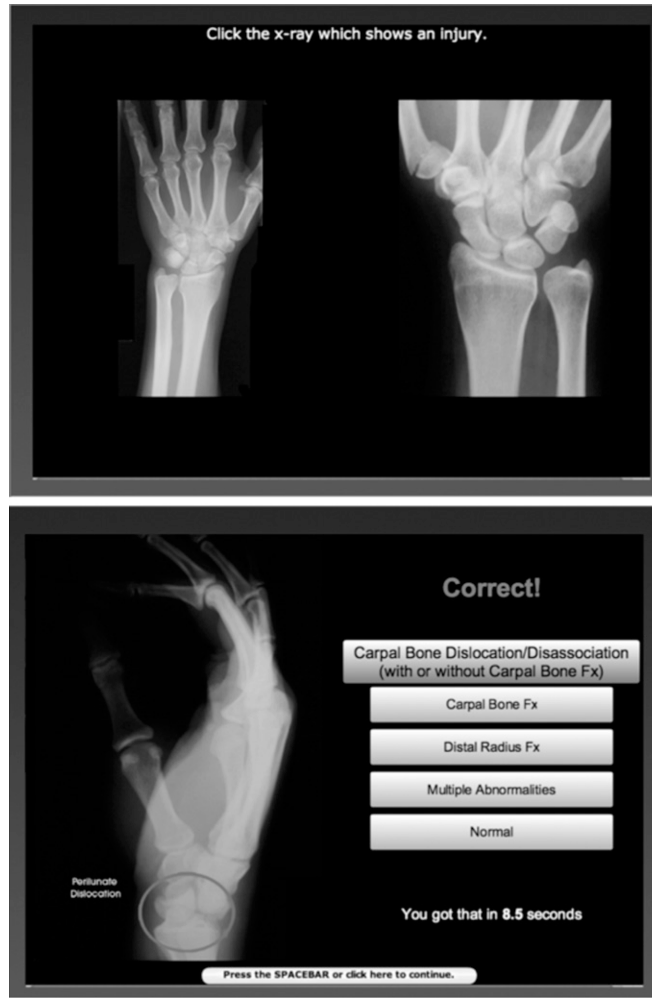

FIGURE 1. Examples of Some Trial Types in the Wrist X-ray PLM.

Perceptual/Adaptive Learning Modules (PALMs) in Dermatology and Histopathology

In collaboration with the David Geffen UCLA School of Medicine, we have recently developed and tested two computer-based PALMs in the pre-clerkship curriculum for first- and second-year medical students, one for recognizing pathologic processes in skin histology images (Histopathology PALM) and the other for identifying skin-lesion morphologies (Dermatology PALM). The goal was to assess their ability to develop pattern recognition and discrimination skills leading to accuracy and fluency in diagnosing new instances of disease-related patterns. We used pre- and post-test design, with each test consisting of the presentation of a visual display along with possible answers for categorizing it. No feedback was given in the assessments. The PALM, given to UCLA medical students in between pre- and post-test, consisted of short interactive learning trials requiring the learner to classify images. The PL components in these modules included deploying a large display set, such that instances of categories did not repeat; moreover, as much as possible, irrelevant variables were balanced across categories (For instance, different dermatological conditions involved approximately the same range of body parts in the displays.) The initial PALMs in these domains were simple; they included only a single type of trial (display presentation with verbal category labels). More varied and complex trial types are known to facilitate PL, but these will be explored in subsequent work.

The adaptive learning components included use of category sequencing algorithms, which optimized spacing based on individual performance, as well as implementation of mastery criteria for each category, based on both sustained accuracy and fluency criteria.

The Dermatology PALM, designed to enhance the skin-lesion morphology curriculum presented in Year 2, consisted of 12 categories of lesion morphologies and was completed by 161 of the 162 second-year students. The Histopathology PALM was designed to complement the skin histopathology curriculum of Year 1 students by enhancing their ability to discriminate the different patterns of presentation observed for cell and tissue injury/repair, inflammation, neoplasia, and normal skin histology images, each at high-power and low-power magnifications. This module was completed by all 161 first-year students. The Histopathology PALM was also required of Year 2 students both to measure retention of the subject from Year 1 and to serve as review and enhanced learning of the material. The Dermatology PALM was offered to Year 1 students, as a control, on a voluntary basis and was completed by 78 students. These modules were completed quickly, with learning criteria typically reached in 15–35 minutes.

As shown in Table II, substantial improvements between pre- and post-test scores were observed, with large (mean effect sizes >0.7) and highly significant ($p < 0.0001$) increases in accuracy and speed in categorizing previously unseen images. Comparing performances for Years 1 and 2 on each of the modules, it can be seen that pre-test scores were much higher for dermatology lesion morphology in Year 2 than in Year 1, which is expected because the students in Year 2 had recently received lectures and an online learning experience. In contrast, this material was touched on only briefly for Year 1 students. Post-test scores, however, were highly similar. Histopathology pre- and post-test scores were similar for Year 1 and 2 students (Table II), showing strong learning gains for both groups. Finally, students reported that the PALMs increased their confidence and were useful, and they indicated that they would like more of these in other units.

APPLICATIONS OF PERCEPTUAL AND ADAPTIVE LEARNING TECHNOLOGIES IN MEDICAL SIMULATION

Although efforts are in their infancy, the promise of perceptual and adaptive learning technologies for improving medical learning is already obvious. Not much work, however, has yet addressed procedural learning and simulation. These areas are ripe for development, as these new technologies are well suited to getting the most from simulation training. Simply having cutting-edge simulations does not solve the problem of how to improve learning. Perceptual–procedural learning technologies and adaptive methods using objective criteria of learning have much to offer in this regard. In this section we note some issues, benefits, and considerations in applying these new technologies to simulation.

Perception–Action Loops in Procedural Learning

We often think of skilled practitioners, such as pilots or surgeons, as having “good hands,” but the key to their skills

TABLE II. Results of Dermatology and Histopathology PALMs with First and Second Year Medical Students

	Pre-Test[a]	Post-Test[a]	*p*	*t*(*df*)	Effect Size	*N*
Year 1 Histopath						
Accuracy	54% (13%)	66% (12%)	<0.0001	9.6	0.98	161
RT	15.82 (11.20)	6.16 (2.33)	<0.0001	11.6	1.19	161
Year 1 Derm (optional)						
Accuracy	66% (11.5%)	84% (8.1%)	<0.0001	12.5	1.55	79
RT	8.36 (6.11)	3.90 (1.03)	<0.0001	7.0	1.00	79
Year 2 Histopath						
Accuracy	51% (13.5%)	64% (14.4%)	<0.0001	7.9	0.90	162
RT	8.87 (4.6)	5.30 (2.7)	<0.0001	11.1	0.95	161[b]
Year 2 Derm						
Accuracy	82% (8.9%)	88% (8.9%)	<0.0001	7.62	0.76	161
RT	9.01 (4.77)	3.72 (094)	<0.0001	14.5	1.54	161

Source: Krasne, Hillman, Rimoin, Burke, Kim, Drake, Craft and Kellman (unpublished data). [a]Mean percent correct is shown in each assessment, with the standard deviation (SD) shown in parentheses. [b]One outlier (pre-test value 182 seconds) was excluded from the analysis.

is often the expert pickup of the information that guides the procedure (cf., Kellman and Kaiser[13]). A surgeon, for example, must recognize anatomy in novel cases, distinguish various tissues and structures, and sense the position, progress, and force of instruments. Surgery involves a host of delicate perceptual–procedural learning tasks. In such procedural tasks, much of the learning consists of improvements in the pickup of information. PL technology can be readily adapted to perception–action loops, in which perceptual discriminations and perceptually guided actions are objectively assessed and accelerated through adaptive spacing and sequencing methods.

Realism in Perceptual–Procedural Learning

High-quality simulations offer ideal synergies with perceptual–procedural learning technology. Because PL requires becoming attuned to subtle patterns of information, richer simulation improves the likelihood that training will be relevant to actual practice. Simulation also allows the variation in displays and scenarios that is crucial for allowing perceptual–procedural learning processes to distinguish crucial invariances from irrelevant variation and discover the best information for guiding action.

Objective Scoring, Assessment, and Certification

A truly revolutionary opportunity afforded by realistic simulation for procedure training, combined with adaptive methods, is the opportunity to objectively assess performance and use performance data to optimize the learning and implement mastery criteria. With suitable tracking technology, a learner's movements in a virtual space may be tracked and scored for accuracy and speed in a variety of tasks. Use of the ARTS system offers particular advantages, in that fluency, as well as accuracy, is used both in optimizing the learning process and also as a *goal* of learning. It has been said that surgeons fall into three categories: fast and good, fast and bad, and slow and bad; there are no slow, good surgeons. Although speed is known to be crucial for effective execution of medical procedures, it has not typically been used systematically in training or assessment. Obtaining and using the objective data required to unlock the benefits of perceptual and adaptive learning techniques may require considerable investment, but doing so will produce accelerated training, reliable certification, practice and refreshment of seldom-used skills in refresher training, and reduction of errors in subsequent medical practice.

CONCLUSION

For effective learning, cutting-edge simulation must be combined with cutting-edge learning techniques. Both in simulation and in many other areas of medical learning, two broad-based innovations offer remarkable potential to accelerate learning and enhance performance. Perceptual–procedural learning technology offers ways of bringing intuitive pattern recognition, interpretation of new cases, fluent processing, and procedural execution into the realm of systematic and objective instruction. Adaptive learning technology, such as the ARTS system, improves learning by attuning the level, spacing, and sequencing of learning events to each individual learner, allowing more efficient learning, better retention, and certification of mastery. These technologies apply to many medical domains, and it will be exciting to turn their potential into reality and see the benefits both in improved training and medical practice.

ACKNOWLEDGMENTS

I thank Cheryl Hein, Carmack Holmes, Erik Dutson, Robert Bjork, Christine Massey, K.P. Thai, and Everett Mettler for helpful discussions and contributions.The author gratefully acknowledges support from: the National Science Foundation (NSF) REESE Program award 1109228; the National Institutes of Health (NIH) award 5RC1HD063338; the U.S. Department of Education, Institute of Education Sciences (IES), Cognition and Student Learning Program (CASL) awards R305A120288 and R305H060070; the UCLA Center for Advanced Surgical and Interventional Technology (CASIT); and the Institute for Innovative Technology in Medical Education (iInTIME); as well as a U.S. Department of Education, IES SBIR award ED-IES-10-C-0024 to Insight Learning Technology, Inc. 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. Bereiter C, Scardamalia M: Rethinking learning. In: The Handbook of Education and Human Development, pp 485–513. Edited by Olson DR, Torrance N. Malden, MA, Blackwell Publishers, 1998.
2. Kellman PJ: Perceptual learning. In: Stevens' Handbook of Experimental Psychology, Vol. 3, Learning, Motivation, and Emotion, Ed 3, pp 259–99. Edited by Gallistel R. New York, Wiley, 2002.
3. Gibson EJ: Principles of Perceptual Learning and Development. New York, Prentice-Hall, 1969.
4. Chase WG, Simon HA: Perception in chess. Cogn Psychol 1973; 4(1): 55–81.
5. Goldstone R: Unitization during category learning. J Exp Psychol 2000; 26(1): 86–112.
6. Goldstone RL: Perceptual learning. Annu Rev Psychol 1998; 49: 585–612.
7. Schneider W, Shiffrin RM: Controlled and automatic human information processing: I. Detection, search, and attention. Psychol Rev 1977; 84(1): 1–66.
8. Ericcson KA, Smith J: Toward a General Theory of Expertise. Cambridge, UK, Cambridge University Press, 1991.
9. Fahle M, Poggio T (editors.): Perceptual Learning. Cambridge, MA, MIT Press, 2002.
10. Kellman PJ, Garrigan PB: Perceptual learning and human expertise. Phys Life Rev 2009; 6(2): 53–84.
11. de Groot A: Thought and Choice in Chess. The Hague, Mouton, 1978. (Original work published 1946.)
12. Bolstad CA, Endsley MR, Cuevas HM: A theoretically-based approach to cognitive readiness and situation awareness assessment. In: Teaching and Measuring Cognitive Readiness. Edited by O'Neil HF, Perez RS, Baker EL. New York, NY, Elsevier.
13. Kellman PJ, Kaiser MK: Perceptual learning modules in flight training. In: Proceedings of the 38th Annual Meeting of the Human Factors and Ergonomics Society, 1994, pp 1183–7, Nashville, TN, Human Factors and Ergonomics Society.
14. Kellman PJ, Massey CM, Son J: Perceptual learning modules in mathematics: enhancing students' pattern recognition, structure extraction, and fluency. Top Cogn Sci 2010; 2(2): 285–305.

15. Goldstone RL, Landy D: How abstract is symbolic thought? J Exp Psychol Learn Mem Cogn 2007; 33(4): 720–33.
16. Massey CM, Kellman PJ, Roth Z, Burke T: Perceptual learning and adaptive learning technology: developing new approaches to mathematics learning in the classroom. In: Developmental and Learning Sciences Go to School: Implications for Education. Edited by Stein NL. New York, NY, Taylor & Francis, 2010.
17. Wise JA, Kubose T, Chang N, Russell A, Kellman PJ: Perceptual learning modules in mathematics and science instruction. In: Teaching and Learning in a Network World: TechEd 2000 Proceedings, pp 169–76. Edited by Hoffman P, Lemke D. Amsterdam, IOS Press, 2000.
18. Goldstone RL, Landy D, Son JY: A well-grounded education: the role of perception in science and mathematics. In: Symbols, Embodiment, and Meaning. Edited by de Vega M, Glenberg A, Graesser A, pp 327–55. New York, Oxford University Press, 2008.
19. Kellman PJ, Massey CM, Roth Z, et al: Perceptual learning and the technology of expertise: fraction learning and algebra. Cogn Prag 2008; 16(2): 356–405 (special issue).
20. Silva AB, Kellman PJ: Perceptual learning in mathematics: the algebra-geometry connection. In: Proceedings of the Twenty-First Annual Conference of the Cognitive Science Society, pp 683–8. Edited by Hahn M, Stoness SC. Mahwah, NJ, Erlbaum, 1999.
21. Atkinson RC: Computerized instruction and the learning process. Am Psychol 1968; 23(4): 225–39.
22. Atkinson RC: Optimizing the learning of a second-language vocabulary. J Exp Psychol 1972; 96: 124–9.
23. Pavlik PI Jr., Anderson JR: Using a model to compute the optimal schedule of practice. J Exp Psychol Appl 2008; 14(2): 101–17.
24. Mettler E, Massey C, Kellman PJ: Improving adaptive learning technology through the use of response times. In: Proceedings of the 33rd Annual Conference of the Cognitive Science Society. Edited by Carlson L, Holscher C, Shipley T. Boston, MA, Cognitive Science Society, 2011.
25. Pyc M, Rawson K: Examining the efficiency of schedules of distributed retrieval practice. Mem Cognit 2007; 35: 1917–27.
26. Benjamin AS, Bjork RA: Retrieval fluency as a metacognitive index. In: Implicit Memory and Metacognition: The 27th Carnegie Symposium on Cognition. Edited by Reder LM, pp 309–38. Hillsdale, NJ, Erlbaum, 1996.
27. Karpicke J, Roediger HL: Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention. J Exp Psychol Learn Mem Cogn 2007; 33: 704–19.
28. Pyc M, Rawson K: Testing the retrieval effort hypothesis: does greater difficulty correctly recalling information lead to higher levels of memory? J Mem Lang 2009; 60: 437–47.
29. Storm BC, Bjork RA, Storm JC: Optimizing retrieval as a learning event: when and why expanding retrieval practice enhances long-term retention. Mem Cognit 2010; 38: 244–53.
30. Balota DA, Duchek JM, Logan JM: Is expanded retrieval practice a superior form of spaced retrieval? A critical review of the extant literature. In: The Foundations of Remembering: Essays in Honor of Henry L. Roediger III, pp 83–106. Edited by Nairne JS. New York, Psychological Press, 2007.
31. Cull WL: Untangling the benefits of multiple study opportunities and repeated testing for cued recall. Appl Cogn Psychol 2000; 14: 215–35.
32. Landauer TK, Bjork RA: Optimum rehearsal patterns and name learning. In: Practical Aspects of Memory, pp 625–32. Edited by Gruneberg MM, Morris PE, Sykes RN. London, UK, Academic Press, 1978.
33. Mettler E, Kellman PJ: Adaptive sequencing in perceptual learning. J Vis 2010; 10(7): 1098.
34. Kellman P.: ACAT Technology for Optimal Sick Call Performance, Technical Report, US Army RDECOM, Contract W91CRB-09-C-0093, 2011.