Accelerating expertise: Perceptual and adaptive learning technology in medical learning

Philip J. Kellman & Sally Krasne

To cite this article: Philip J. Kellman & Sally Krasne (2018) Accelerating expertise: Perceptual and adaptive learning technology in medical learning, Medical Teacher, 40:8, 797-802, DOI: 10.1080/0142159X.2018.1484897

To link to this article: https://doi.org/10.1080/0142159X.2018.1484897

Published online: 09 Aug 2018.

Submit your article to this journal

Article views: 342

View Crossmark data
Accelerating expertise: Perceptual and adaptive learning technology in medical learning

Philip J. Kellmana and Sally Krasneb

aDepartment of Psychology, University of California, Los Angeles, CA, USA; bDepartment of Physiology, David Geffen School of Medicine, University of California, Los Angeles, CA, USA

ABSTRACT
Rationale: Recent advances in the learning sciences offer remarkable potential for improving medical learning and performance. Difficult to teach pattern recognition skills can be systematically accelerated using techniques of perceptual learning (PL). The effectiveness of PL interventions is amplified when they are combined with adaptive learning (AL) technology in perceptual–adaptive learning modules (PALMs).

Innovation: Specifically, PALMs incorporate the Adaptive Response Time-based Sequencing (ARTS) system, which leverages learner performance (accuracy and speed) in interactive learning episodes to guide the course of factual, perceptual, or procedural learning, optimize spacing, and lead learners to comprehensive mastery. Here we describe elements and scientific foundations of PL and its embodiment in learning technology. We also consider evidence that AL systems utilizing both accuracy and speed enhance learning efficiency and provide a unified account and potential optimization of spacing effects in learning, as well as supporting accuracy, transfer, and fluency as goals of learning.

Results: To illustrate this process, we review some results of earlier PALMs and present new data from a PALM designed to accelerate and improve diagnosis in electrocardiography.

Conclusions: Through relatively short training interventions, PALMs produce large and durable improvements in trainees’ abilities to accurately and fluently interpret clinical signs and tests, helping to bridge the gap between novice and expert clinicians.

Introduction
Recent advances in the learning sciences offer profound potential to improve medical education. In this paper, we describe two areas of recent innovation that offer new principles and new learning technology in medical learning. The first, perceptual learning (PL) approaches, teach pattern recognition, structural intuition, and fluency. The second, adaptive learning (AL) technologies, optimize learning for each individual, embed objective assessment throughout learning, and implement objective mastery criteria. We also describe recent combinations of these in perceptual–adaptive learning modules (PALMs), summarizing their effects in medical learning domains and providing a detailed example of their formulation and outcomes based on a PALM for training interpretation of electrocardiograms.

Conceptions of learning
Underlying much of our work are changing conceptions of what learning is. As is the case in most instructional settings, medical learning is dominated by declarative knowledge – facts and concepts that can be verbalized, and procedural knowledge – sets of steps that can be enacted. These are surely important parts of learning; however, they are neither exhaustive nor do they cover much of what a medical student or resident needs to master in order to be an effective practitioner. Bereiter and Scardamalia (1998) suggested that a pervasive “folk psychology” stereotype about what learning is affects ordinary people, practitioners, and learning researchers alike. They called this implicit standard view the “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.

This conception is much too narrow, and what is missing relates to persistent problems, and considerable frustration, in learning and instruction. Students who have been carefully taught and who have diligently absorbed declarative and procedural inputs fail to recognize key structures and patterns in real-world tasks, such as interpreting radiographs, ECGs, cytology, and other clinical images and tests. Trainees 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, causing them to be impaired in demanding, complex, or time-limited tasks. In the realm of medicine, there is clearly a gap between the foundational knowledge gained in medical school and the ability to recognize relevant, clinical patterns during residency and beyond.

In the literature on expertise, rather than learning, we find important clues to what is missing. Studies of experts in any domain reveal that they extract and incorporate information differently from novices (Chase and Simon 1973; Kellman and Garrigan 2009). In particular, experts selectively pick up meaningful structures and relations while ignoring irrelevancies, and they process task-relevant information rapidly and with low attentional load. Much of their expertise arises from perceptual systems that have become progressively attuned and adapted to the structure of information in the task domain.

**Perceptual learning**

How do these expert abilities arise? They are products of perceptual learning (PL). PL is broadly defined as experience-induced improvements in the extraction of information (Gibson 1969). For example, one learns to recognize the voices of family and friends, and distinguish among them, based on experience rather than an analysis of the frequency characteristics of each. In analogous fashion, experienced physicians are able to accurately and efficiently pick out key features of patient scripts and interpret patterns in clinical test results based on experience, rather than by recalling facts and procedures that they were initially taught as guides to making such interpretations. A wealth of research supports the notion that, with appropriate practice in any domain, the brain progressively improves information extraction to optimize task performance in that domain (for reviews, see Gibson 1969; Goldstone 1998; Kellman 2002; Kellman and Garrigan 2009).

PL effects improve information extraction in a variety of ways. Kellman (2002) argued that there are two broad categories of improvements: discovery and fluency effects. Discovery effects involve finding the information that is relevant to a domain or classification. Fundamental among discovery effects is selection (Gibson 1969; Petrov et al. 2005): We discover and extract the information relevant for a task, ignoring or inhibiting information that is irrelevant. We come to process complex relationships in the available input to which we were initially insensitive – an improvement in sensitivity in a signal detection sense. PL in the contemporary sense involves improved use of information available in the stimulus environment rather than changing criterion or bias (Gibson and Gibson 1955; Kellman and Garrigan 2009). Fluency effects involve the efficiency of extracting discovered information – faster encoding, pickup of larger chunks (Chase and Simon 1973; Goldstone 2000) or more parallel processing and reduced cognitive load (Schneider and Shiffrin 1977). Discovery and fluency may work iteratively in that a dividend of more fluent performance is that it frees up resources for discovery of even higher-order task-relevant information (Bryan and Harter 1899; Kellman and Garrigan 2009).

These effects are evident in experts in many areas of medical practice. As has been documented in a number of other domains (Kellman and Garrigan 2009), experts in medical image interpretation locate targets much more quickly and accurately than novices and use more efficient search patterns (Krupinski 2010). As is typically the case, PL is highly domain-specific (Kellman and Garrigan 2009), and expertise in medical image interpretation in a given area is specifically related to repeated experience with relevant images (e.g., radiologists are not better than lay people at detecting non-medical targets, such as finding Waldo in "Where’s Waldo" picture books; Nodine and Krupinski 1998). Interestingly, while expert performance indicates remarkable domain-specific changes in sensitivity to relevant information, it is not reliably accompanied by conscious awareness of how the detection or classification is being accomplished.

At higher levels of pattern recognition, a surgeon recognizes anatomy in novel cases, distinguishes various tissues, structures, and planes, and senses the position, progress, and force of instruments; emergency medical doctors interpret patterns on monitors in trauma care, and experienced diagnosticians more rapidly and accurately see relations among tests and symptoms, and combine information from different sources to make accurate diagnoses.

**Perceptual learning technology**

Conventional declarative and procedural instruction does little to advance expert pattern recognition and fluency. In most domains, in fact, there has been a tacit assumption that we cannot teach this kind of knowing. Among the problems in addressing PL with conventional instructional methods is that much of PL occurs unconsciously (c.f. Reber 1993; Seitz and Watanabe 2003; Mettler and Kellman 2006). In accord with this assumption, radiologists, surgeons, and pathologists, as well as chemists, pilots, and air traffic controllers, advance through apprenticeship: The eye of the expert is thought to emerge from “seasoning,” or “experience.”

PL grows from many classification episodes and feedbacks from encountering sufficient variation within and between categories to be learned (Kellman and Garrigan 2009). Computational models of PL stress the discovery and selective weighting of relevant features and relations (Petrov et al. 2005; Kellman and Garrigan 2009), a process that often occurs gradually across many classification episodes.

Understanding that pattern recognition learning grows by classification events opens the possibility of systematically addressing and accelerating PL using appropriate computer-based interventions. We have developed an emerging technology of PL (Kellman and Kaiser 1994; Kellman 2013; Kellman and Massey 2013; Mettler and Kellman 2014), and in recent research, it has been successfully applied to a number of medical learning domains (Krasne et al. 2013; Rimoin et al. 2015; Thai et al. 2015; Romito et al. 2016; Krasne et al. under review). PL is systematically advanced by presenting learners with many short, interactive episodes during which they encounter a sufficient number and variety of exemplars, which they classify into appropriate categories, to train both accurate
generalization to new exemplars of the same category (e.g. learning to recognize related examples despite their wide variation in appearance) and differentiation (learning to make fine discriminations between easily confusable categories, such as between melanomas and benign moles or seborrheic keratoses). Items in the software are organized into target categories (e.g. diagnostic categories, structural identifications). Item sets are large, so that individual exemplars are unlikely to repeat, and learners master each learning category to proficiency, with a continuous stream of specific feedback for both correct responses and various error types.

Adaptive learning technology

Medical learning could be vastly improved by technology that adapts to the needs of the individual learner. Students have different starting points, receive instruction of varying quality, and differ in components of instruction that they learn well or poorly. Testing often occurs 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. Improved systems would use accuracy and fluency measures to guide the spacing (how soon in a sequence a category should be repeated) and sequencing of learning events. Likewise, performance measures would guide the learner to objective mastery criteria for all components of learning tasks. These are benefits realizable from recent innovations in AL technology.

There have been a variety of efforts in AL, and evidence is strong that they produce robust improvements in learning (e.g. Atkinson 1972; Pavlik and Anderson 2008). Some limitations of most approaches are as follows: (1) They use elaborate models that require obtaining prior data from relevant learners and subject matter to estimate parameters (e.g. Atkinson 1972; Pavlik and Anderson 2008); (2) they are primarily focused on accuracy data alone; and 3) they either do not incorporate spacing effects in learning (Atkinson 1972) or they add spacing elements in an ad hoc manner (Pavlik and Anderson 2008).

Considerable research indicates the importance of spacing in optimizing learning and retention. But what is optimal spacing? Some evidence suggests that learning is best using a fixed schedule of expanded presentation intervals (Landauer and Bjork 1978; Storm et al. 2010). Other work suggests equal spacing intervals produce the best retention (Karpicke and Roediger 2007). Recent work (Mettler et al. 2016) suggests that there is no single, correct answer to the question of what predetermined recurrence schedule optimizes learning. The most ideal time for a memory item or a category in perceptual classification to recur is when the learner can still respond successfully with some effort (Pyc and Rawson 2009; Mettler et al. 2016), but this interval depends on current learning strength of an individual learner for that item or category. Any predetermined schedule is non-adaptive and thus is insensitive to differences among learners, differences among items, and interactions of the two. Further, spacing paradigms based solely on accuracy as a measure of learning are unable to distinguish between slower deliberative processes versus automatic pattern recognition or rapid memory retrieval. Rapid performance with low attentional load is important to competent performance in complex and/or time-critical tasks, such as driving, surgical procedures, or decision-making in medical emergency settings. Adding response time (RT) provides a window into the type of processing the learner is using and can also be used to ensure fully fluent performance. Fluent mastery is realized when a learner can respond accurately and rapidly over long delays.

Adaptive Response Time-based Sequencing (ARTS) (Mettler et al. 2011) is a novel approach to AL that incorporates recent research findings regarding spacing and other principles of learning and memory in a natural way and uses both accuracy and response speed in spacing and sequencing categories and for setting learning criteria. In ARTS, response accuracy and the speed of (accurate) responses (fluency) are indicators of current learning strength and serve as inputs to a dynamic spacing algorithm that uses a priority score system to automatically space and interleave active learning categories. Each category is assigned a priority score indicating the relative benefit of a new exemplar of that category appearing on the next learning trial, and all learning categories compete simultaneously as a function of their priority score. Priority scores for each category are updated after every trial as a function of accuracy, response time, and trials elapsed since the previous presentation (Mettler et al. 2011). As learning strength for a given category increases, the ARTS algorithm automatically generates lower priority, and longer recurrence intervals, as an inverse function of the log of RT. Figure 1 illustrates how this sequencing and spacing are determined. The left-hand image illustrates the case of an incorrect answer, for which another exemplar of the category will be presented soon. The middle image illustrates the case in which an item is quickly classified into its correct category, for which there will be a long delay before another exemplar from that category is presented. The right-hand image illustrates the case of a correct, but slow, classification event, for which another exemplar of the category will be presented with an intermediate number of intervening trials.

For category learning, learners must respond accurately and fluently to novel exemplars of categories across delays, indicating that they are picking up key diagnostic

---

**Figure 1.** Illustration of adaptive spacing and sequencing based on response accuracy and speed.
Fluency is used both in arranging the flow of interactive learning events and is itself a goal of learning, included in mastery criteria. Meeting mastery criteria requires successful responses within a designated maximum response time to multiple successive spaced exemplars of a category. For example, a mastery criterion for correctly recognizing a specific histopathologic process (category) might require three consecutive, accurate identifications of exemplars of that process with each accurate response occurring within 10 seconds (s). Published research shows that ARTS offers clear advantages in efficiency and durability of learning in general (Mettler et al. 2011; Mettler et al. 2016) and in PL specifically (Mettler and Kellman 2014). The ARTS outperforms random presentation (Mettler and Kellman 2014) and also outperforms a classic AL system (Atkinson 1972) in tasks involving learning of factual items (Mettler et al. 2011).

A feature that adds considerable power to the system is the use of category retirement. Upon meeting mastery criteria, a category is removed from the learning set, which allows AL to focus each learner’s effort where it is needed most, on those categories which have not yet been mastered. Pyc and Rawson (2007) 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. The assessment capabilities guide learning to criterion and also offer rigorous, objective bases for certification. Furthermore, the ARTS system is easily configured to provide remarkably efficient recurrent training. Rapid, automated assessments determine which categories, classifications, facts, or concepts are still well-learned and which require refreshment. For the latter, learning is resumed and is guided to objective mastery criteria.

Perceptual and adaptive learning modules (PALMs) in medical education

The ARTS system is quite general in that it applies to factual information, procedures, or perceptual classification. Of special interest in medical learning, however, is the combination of AL with PL interventions, since many domains of medical practice involve complex displays (e.g. areas such as radiology, dermatology, pathology, electrocardiography, ultrasound, surgery), that involve extracting key features or pattern recognition. In recent work, we have developed and tested online PALMs in a number of challenging areas of medical learning. These PALMs apply ARTS to perceptual category learning, using categories relevant to interpretations of medical classifications such as clinical tests (e.g. electrocardiograms, pathologic processes, fetal heart rate tracings), identification of anatomical structures (e.g. in CT images and ultrasound recordings), and characterization of lesions (e.g. dermatology). Each category is comprised of a large enough number of exemplars that repetition of a specific exemplar is uncommon. These PALMs have consistently produced remarkable acceleration in learning, which required relatively short interventions and was durable. Details on several of these interventions have been published (Krasne et al. 2013; Rimoin et al. 2015; Thai et al. 2015; Romito et al. 2016).

Integration of ARTS and PL into a PALM and the outcomes one can observe can be best understood from a concrete example, the ECG Morphology PALM (Krasne et al. under review). This PALM aims to train interpretation of 15 categories (diagnoses), 12 of which are manifested as changes in the shapes of traces within 12-lead electrocardiograms (acute or old/indeterminant myocardial infarctions, bundle branch blocks, axis deviations, atrial enlargements, ventricular hypertrophies), the other three being related to heart rate (sinus bradycardia, tachycardia, or a normal sinus rhythm). Within each category, there are typically 30–40 ECGs (i.e. exemplars), all from different patients but having the same diagnostic interpretation. The PALM, itself, unfolds as a sequence of trials, each displaying a 12-lead ECG image along with four answer choices, only one being correct. The trainee is allowed 30 s to choose an answer, after which feedback is given in the form of the correct answer, a text description of the category’s key features, and indicators of these features on the ECG tracing itself. For correct answers, the response time is also shown. Sequencing and spacing of exemplar presentation from each category is determined based on the ARTS priority system with the minimum spacing for category repetition set to three intervening trials. The objective mastery criteria set for retiring each category are three consecutive, accurate answer choices for the category, each within a target response time of 15 s (fluent responses). The PALM effectiveness and durability are assessed via a pretest, posttest, and delayed test. Each test consists of two unique exemplars per category and provides no feedback or adaptive spacing and sequencing.

Figure 2. The ECG Morphology PALM significantly improved accuracy (A) and fluency (B) in 12-lead ECG interpretation. A pretest (black bars) taken by MS3&4 students and R2&3 emergency medicine residents shows a progression in performance with increased educational level. A posttest after ECG PALM training (grey bars) shows all groups reaching approximately the same high level of accuracy and their fluency roughly doubling, $p < 0.0001$ for each pre- and posttest comparison. Improvement persisted at least 1 year following training (C). Error bars are one standard error.
This ECG Morphology PALM was used in training third and fourth year medical students (MS3 and MS4), and second and third year emergency medicine residents (R2 and R3). Figure 2 illustrates the effectiveness and durability of the PALM, with “accuracy” reflecting the percentage of trials correctly identified within the allotted 30-s window per trial, and “fluency” reflecting the subset of accurate answers made within the target response time of 15 s. Although the medical students started at much lower levels than the residents, they reached close to the same levels of accuracy and fluency following PALM training. In addition, a substantial proportion of their improvement was maintained over at least 1 year, and the learning was efficient; training times averaged between 1.5 h (MS3s) and 40 min (residents).

Conclusions
Incorporating approaches that enhance PL and developing a flexible, user-centric approach to sequencing and spacing material to be learned, based on the combination of an individual's accuracy and response time (fluency) are two new approaches to enhance medical training, each based on the knowledge gained from research in cognitive science. Their combination in the form of PALMs, along with the ability to set competency requirements for determining when a learning category has been sufficiently mastered, can provide a pathway for training each individual up to a desired level of proficiency and can serve to maintain that level of competency as well. Recent reports by others have also recognized the value and potential of combining PL and AL. Evered (in press) reviewed the basic concepts and scientific bases of PALMs and argued that the use of PALM technology would improve training in cytology. Following the publication of data from tests of a PALM in transesophageal echocardiography (Romito et al. 2016), an unsolicited editorial in the British Journal of Anaesthesia commented that perceptual–adaptive learning in PALMs has “…the potential to revolutionize our traditional approaches to learning in anesthesia” (Weller 2016). Based on our own experiences in using PALMs to train medical students and residents in a variety of areas ranging from characterizing skin lesions and discriminating histopathologic processes to categorizing fetal heart rate tracings, interpreting electrocardiograms, and classifying heart functions based on echocardiograms, we, too, think that these tools can have a significant impact in medical education.

Acknowledgments
We gratefully acknowledge support from the agencies listed in the Funding section below. Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of NSF, the U.S. Department of Education, NIH or other agencies. We thank Rachel Oldier for general assistance and Christine Massey for helpful discussions. A number of features of the perceptual and adaptive learning technologies described here are covered by U.S. Patents 7052277 and 9299265. For more information, please contact Info@insightlt.com

Disclosure statement
The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

Funding
This work was supported by the US National Science Foundation [grant number DRL 1644916]; the U.S. Department of Education, Institute of Education Sciences (IES), Cognition and Student Learning (CASL) Program [grant number R305A120288]; the US National Institute of Health [grant number 5RC1HD063338], and a grant from iNtTIME/UMed.

Notes on contributors
Philip J. Kellman, PhD, is Distinguished Professor of Psychology at UCLA and Adjunct Professor in the Department of Surgery at the David Geffen School of Medicine, UCLA. Sally Krasne, PhD, is Research Professor of Physiology at the David Geffen School of Medicine, UCLA and a member of UCLA’s Brain Research Institute.

References
Evered A. (in press). Perceptual and adaptive learning modules (PALMs) and their potential to transform cytology training programmes. Cytopathology. DOI:10.1111/cyt.12578


