Deliberate Practice and Acquisition of Expert Performance: A General Overview

K. Anders Ericsson, PhD

Abstract

Traditionally, professional expertise has been judged by length of experience, reputation, and perceived mastery of knowledge and skill. Unfortunately, recent research demonstrates only a weak relationship between these indicators of expertise and actual, observed performance. In fact, observed performance does not necessarily correlate with greater professional experience. Expert performance can, however, be traced to active engagement in deliberate practice (DP), where training (often designed and arranged by their teachers and coaches) is focused on improving particular tasks. DP also involves the provision of immediate feedback, time for problem-solving and evaluation, and opportunities for repeated performance to refine behavior. In this article, we draw upon the principles of DP established in other domains, such as chess, music, typing, and sports to provide insight into developing expert performance in medicine.

E ducation and training have a very long history. Ever since Greek civilization, there has been an important distinction made between general education and vocational training, which requires the acquisition of a high level of skill. Education, as proposed by Plato and Aristotle, discouraged specialization and focused on developing thinking skills and general knowledge. With respect to skilled manual work, primarily performed by slaves in Athens,1 Plato proposed an early start of training and restricted engagement to a single craft.2,3 When discussing medical expertise, Plato contrasted the routine application of medical procedures by a slave doctor to the thoughtful diagnosis and reasoning about treatment and explanation to the patient by an expert doctor.4 Competency among doctors was thus thought to be best assessed by examining their superior knowledge5 and ability to teach others.6 The primary additional factor to achieve competence was believed to be due to accumulated experience, and thus age was expected to be correlated with wisdom. Plato argued that the ideal doctor was one with significant experience—ideally, personal experiences of recovering from many diseases.7

Since Aristotle, the issues of how to evaluate and certify expertise and the tension between theoretical understanding and practical skill have remained. The level of expertise of practitioners has been monitored and credentialed by guilds and more recently by professional organizations.8 In the United States, students typically require a general education, such as a college education, before they can be admitted to study at professional schools. Following primarily theoretical training at professional schools, graduates are trained as apprentices to experienced practitioners until they earn the credentials to practice independently.

Consistent with the distinction between general theoretical knowledge and professional skill, early traditional models of skill and expertise9–13 distinguish different stages of development. The first stage (novice) involves following the teachers’ instruction and applying rules and procedures step by step. With increasing experience, the student becomes more able to generate the same outcomes faster and more efficiently. After extensive experience, individuals become experts and are able to respond rapidly and intuitively. Some domains, such as driving a car, are simple and “almost all novices [beginners] can eventually reach the level we call expert.”11 In other more complex domains, such as telegraphy and chess, it may take decades to reach the highest levels.13,14 Some researchers even explicitly reject the idea that expert behavior and performance needs to be uniformly superior to less experienced individuals.11 The pioneering research on expertise13 emphasized improvements in performance due to
experience in the domain. In studies of medical doctors and nurses, it was typical to search for experts by using peer-nomination procedures among highly experienced professionals.\(^9,15,16\)

In the 1980s the definition of expertise based on accumulated knowledge, extensive professional experience, and peer nominations was becoming increasingly criticized. Numerous empirical examples were reported where “experts” with extensive experience and extended education were unable to make better decisions than their less skilled peers or even sometimes than their secretaries.\(^17\) Early studies were unable to establish superior accuracy of the peer-nominated best general physicians, when compared to a group of undistinguished physicians.\(^15,16\) Similar findings were subsequently attained for clinical psychotherapists, where more advanced training and longer professional experience were unrelated to the quality and efficiency of treatment outcomes.\(^18\) In addition, examinations of the cognitive mechanisms that mediated the actions of individuals exhibiting consistently superior performance revealed a complex structure that could not be accounted for by a mere accumulation of experience and knowledge.\(^19\) In response to these criticisms, Ericsson and Smith\(^19\) proposed the redirection of research from studying the behavior of socially recognized experts toward the study of reproducibly superior performance in a given domain.

## The Scientific Study of Expert Performance and Its Acquisition

Establishing a science of superior performance starts with the accumulation of a body of reproducible empirical phenomena.\(^20\) Methods for reproducing the superior performance under standardized and experimental conditions are thus needed. When superior performance can be replicated in the laboratory, its structure can be examined and analyzed with process analysis and experimental methods to reveal the mechanisms underlying superior performance.

Every day, professionals make superior decisions; soccer players make the deciding goal, and scientists make amazing discoveries. However, it is impossible to know if a singular action was due to unique circumstances (if placed in that same context many people would have produced a similar outcome), luck, or a stable superior ability to handle such situations. The Greeks were particularly interested in determining aspects of soldiers’ performance, such as who is the strongest or the fastest and can throw a javelin or discuss the longest. The Greeks developed athletic competitions with standardized events. For example, rather than having athletes run from one point to another in natural terrain, they built straight, flat tracks that were indistinguishable for all runners and devised methods to force runners to start at the exact same time and cross the same finishing line. More recently, these have emerged in competitions in music, dance, and chess that have objective performance measures to identify the winners. In all of these traditional domains, elite individuals reliably outperform less accomplished individuals. There are many groups of professionals who, on a daily basis, perform similar tasks. For example, professional investors have equal access to investment opportunities in the stock market, medical professionals treat patients with similar symptoms, and psychotherapists treat patients with similar reported problems.

An expert athlete, musician, chess master, nurse, or medical doctor is expected to be able to respond to emerging task demands. It is part of the definition of an expert performer that they are able to perform at virtually any time with relatively limited preparation. Elite athletes must be ready to compete, even if the competition is delayed for a few hours, or even a day, because of bad weather. Similarly, an emergency physician (EP) is expected to be able and ready to help when they encounter a distressed patient in the clinic or even in public.

Ericsson and Smith\(^19\) proposed how naturally occurring events can be used to capture the essence of expertise in a given domain. For example, it would be possible to re-create an actual clinical situation by filming a scenario from the responsible doctor’s perspective and decision-making process. This film could then be presented to individual experts and less experienced doctors, the film could be stopped at critical decision points, and the observing doctor may be asked to provide appropriate direction to personnel under normal time constraints. By selecting scenarios where there is only one or a couple of appropriate actions that are not detrimental to the patient, it is possible to evaluate the doctor’s actions and decisions and their ability to assess the essential factors in the time-constrained situation. The pioneering work of selecting critical events was introduced by de Groot\(^21\) in the domain of chess. De Groot extracted critical situations in games between chess masters and then set up a controlled laboratory situation where chess players were sequentially presented with the associated positions (see Figure 1). In another example, since expertise in typing should generalize to any material, all typists may be provided with the same text to type as fast and accurately as possible. The final example given in Figure 1 illustrates the skill of sight reading, where an accompanist is presented with a sheet of music and asked to accompany a singer without having a chance to prepare in advance. The ability of accurately playing as many notes as possible from the music score is what differentiates skilled accompanists from other pianists. Over the past few decades, standardized test situations, where performance can be assessed within an hour, have been developed that are highly correlated to real-world performance, such as tournament performance in chess, golf, and Scrabble; performance in music competitions; and medical diagnosis.\(^22–25\) These findings are consistent with the hypothesis that there is an underlying factor of attained expertise in a domain, where the majority of the tasks can be ordered on a continuum of difficulty.\(^26\) In many domains of expertise, there is a rank ordering of difficulty for mastery of different tasks. For example, dives are given a difficulty score, and music pieces are rated in the number of years of study required before mastery. Similarly, gymnastics, martial arts, mathematics, and many of the sciences have a clear progression of levels defined by mastery of...
increasingly difficult tasks. In sum, a collection of representative tasks that capture the essence of expertise in a given domain may be identified and administered to all participants under controlled and standardized conditions to objectively measure performance.

THE ACQUISITION OF SUPERIOR REPRODUCIBLE (EXPERT) PERFORMANCE

Once we are able to measure individual performance, it is possible to measure the time course of improvement and identify several characteristics that generalize across different domains of expertise. In some domains there is no demonstrable improvement in performance as a function of years of professional experience after completed training. For example, the accuracy of diagnosis of heart sounds and many types of measurable activities of nurses and general physicians do not improve as a function of professional experience, and sometimes the performance even gradually decreases after graduation. In contrast, many traditional domains of expertise, such as arts and sciences, games, and sports, demonstrate improvements that continue for decades.

Based on an analysis of many different domains, consistent patterns of performance level over time (Figure 2) have been observed. When the same uniform adult performance standards are used, abrupt improvements in performance do not occur, and changes over time are gradual. In addition, the age at which experts typically reach their peak performance is in the third and fourth decades for the arts and sciences and somewhat earlier for vigorous sports. Finally, all performers, even the most “talented,” need around 10 years of intense involvement before they reach an international level in established sports, sciences, and arts. Most elite individuals take considerably longer to reach that level. The necessity for years and even decades of required engagement in domain-related activities is the most compelling evidence for the crucial role of experience required to attain high levels of performance. Some of the best evidence for the necessity of improved training methods and expanded practice durations comes from historical comparisons. The most dramatic increases in the level of attained performance over historical time are found in sports. In competitions such as the marathon and swimming events, a large number of today’s serious amateurs could easily beat the gold medal winners of the past.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Presented Information</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chess</td>
<td>Select the best chess move for this position</td>
<td></td>
</tr>
<tr>
<td>Typing</td>
<td>Type as much of the presented text as possible within one minute</td>
<td></td>
</tr>
<tr>
<td>Music</td>
<td>Play the same piece of music twice in same manner</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. Three examples of laboratory tasks that capture the consistently superior performance of domain experts in chess, typing, and music (adapted from Ericsson and Lehmann).

Figure 2. Illustration of the gradual increases in expert performance as a function of age, in domains such as chess. The international level, which is attained after more than around 10 years of involvement in the domain, is indicated by the horizontal dashed line (from Ericsson and Lehmann).
MERE EXPERIENCE VERSUS DELIBERATE PRACTICE

To reconcile the virtual absence of a relation between amount of experience and objective performance in many professional domains on the one hand, and the necessity of many years or even decades of full-time engagement in training for reaching high levels of performance, my colleagues and I tried to identify those domain-related activities necessary for improving performance and classified them as deliberate practice (DP).

Based on a review of research on skill acquisition, we identified a set of conditions where practice had been uniformly associated with improved performance. Significant improvements in performance were realized when individuals were 1) given a task with a well-defined goal, 2) motivated to improve, 3) provided with feedback, and 4) provided with ample opportunities for repetition and gradual refinements of their performance. Deliberate efforts to improve one’s performance beyond its current level demands full concentration and often requires problem-solving and better methods of performing the tasks.

When people are introduced to an unfamiliar domain of activity, such as a new job, sport, or game, they frequently encounter situations where they cannot react fast enough or where they are unable to produce functional actions, resulting in obvious failures. Over time, they are able to figure out adequate responses by practice, problem-solving, and trial-and-error or with help from supervisors, teachers, or colleagues. With further experience they become increasingly able to generate rapid adequate actions with less and less effort—consistent with the traditional theories of expertise and skill acquisition, as is illustrated at the lower arm in Figure 3. After some limited training and experience—frequently less than 50 hours for most recreational activities, such as skiing, tennis, and driving a car—an individual’s performance is adapted to the typical situational demands and is increasingly automated, and they lose conscious control over aspects of their behavior and are no longer able to make specific intentional adjustments. For example, people have automated how they tie their shoelaces or how they stand up from sitting in a chair. When performance has reached this level of automaticity and effortless execution, additional experience will not improve the accuracy of behavior nor refine the structure of the mediating mechanisms, and consequently, the amount of accumulated experience will not be related to higher levels of performance.

In direct contrast, aspiring experts continue to improve their performance as a function of more experience because it is coupled with DP. The key challenge for aspiring expert performers is to avoid the arrested development associated with automaticity. These individuals purposefully counteract tendencies toward automaticity by actively setting new goals and higher performance standards, which require them to increase speed, accuracy, and control over their actions, as is shown in the upper arm of Figure 3. The experts deliberately construct and seek out training situations to attain desired goals that exceed their current level of reliable performance.

The clearest cases of DP are found when children or adolescents have recently gotten involved in active participation in sports, singing, or playing a music instrument. Many of the more “talented” will be encouraged by their parents or coaches to start seeing a professional teacher, who can assess their current performance level and design training activities to improve certain aspects of performance. In these cases, it will be obvious that performance has improved from one week to the next if the students are able to reach a higher target performance, which had previously been outside the range of their performance ability.

After years of daily practice, the aspiring expert performers become able to monitor their performance so they can start taking over the evaluative activity of the teacher and coach. They acquire and refine mechanisms that permit increased control, which allow them to monitor performance in representative situations to identify errors as well as improvable aspects. There is compelling evidence for these complex cognitive mechanisms from studies in expert performance. For example, chess masters can select the best move for a chess position. When the chess position is removed, they are able to report their thoughts and also recall the locations of all the pieces on the chess board virtually perfectly. The superior incidental memory of experts for relevant information for representative tasks have been demonstrated in a large number of domains, such as sports, music, ballet, and medicine.

When expert performers are working on appropriately challenging tasks, there is compelling evidence that their actions are cognitively mediated.
Once a professional reaches an acceptable skill level, more experience does not, by itself, lead to improvements. For example, tennis players will not improve their backhand volley in tennis by playing more games. However, a tennis coach can provide opportunities for DP. Initially, the coach can place relatively easy volleys, followed by increasingly unpredictable and difficult ones, and then later, integrate volleys with backhand volleys into regular game contexts. High-fidelity simulators can provide beginning and advanced EPs with opportunities to improve their performance of medical procedures.

How is it possible to improve one’s ability to plan and select the best action in a given situation? Chess players typically solve this problem by studying published games from chess tournaments between the very best players in the world. They play through the games, one move at the time, to select the best move for a given position before they look at the move made by the chess master. If their move matches the corresponding move selected by the masters, then there is nothing obvious to change. If, on the other hand, the chess master’s move differed from their own selection, it would imply that their planning and evaluation overlooked some aspect of the position. By more careful and extended analysis, the aspiring player is generally able to discover the reasons for the chess master’s move. Most importantly, they can then reflect on their thought process during their faulty move selection and examine how they need to change their planning methods when encountering other related situations in future games. Similarly, it should be possible to collect authentic information about medical cases encountered by surgeons, general physicians and EPs. With the benefit of hindsight, it should be possible to recreate the original patient encounter in a time-constrained context and have the aspiring expert doctor make treatment decisions and get immediate feedback by comparing their decisions with those of expert physicians. Extensive research on how speed of performance can be increased through DP has been conducted on typing. The amount of daily typing by college students is essentially unrelated to current typing speed during a test. The key findings are that special training and efforts to increase typing speed are predictive of the typing speed attained. More generally, individuals can improve typing speed by pushing themselves, as long as they can maintain full concentration. Straining to type at a faster speed—typically around 10%–20% faster than their normal speed—facilitates typists to better anticipate the copy, possibly by extending their gaze further ahead. Similarly, athletes who push themselves overload their muscles, which enact gene expressions that lead to physiologic and anatomic changes, which increase their strength and power.

Several studies and reviews report a consistent relationship between the amount and quality of solitary activities meeting the criteria of DP and performance in a wide range of domains of expertise. To reach a level where one can win international competitions, it is estimated that over 10,000 hours of DP have been generated for several domains.

### Professional Experts’ Reliance on Intuition versus DP with Challenging Cases

Expert performers are able to report their thought processes and critical aspects of the encountered situations. This seems to be inconsistent with the claims for experts’ use of intuition: “When things proceed normally, experts don’t solve problems and don’t make decisions, they do what normally works.” Common situations, such as treating an otherwise healthy adult for a cold or flu, can be handled successfully by virtually anyone, and expert performers would not be expected to demonstrate greater treatment success than other medical service providers. According to the expert performance approach, we need to study challenging task situations with reliable individual differences in performance. Challenging task situations have rarely or even never been experienced first hand. In the domain of medicine, this would involve treating patients with poor prognosis, rare diseases, and conditions involving the interaction of several medical problems and medications. For these types of challenging patients, there are professionals who provide treatments with superior outcomes than their colleagues. Some of these professionals’ performance can be linked to special training and practice motivated by increased standards, such as certification in specialties.

It is not immediately obvious how other professionals gain their performance advantage. When people are engaged in professional activities, public performances, and competitions, it is difficult to engage in learning and training because the priority is on performing consistently at a high level. Even if an individual is aware of a mistake or a failed step, it is not possible to stop the ongoing activity during a public music performance. Instead, the professionals need to proceed and make any necessary adjustments to minimize the perceptible effects of the disruption and maximize the chances for a successful overall outcome. There are some medical domains, such as surgery, where mistakes lead to observable consequences that need to be immediately addressed. Interestingly, in surgery there is evidence for improvements in performance as a function of surgeries of a given type.

In a professional environment with real-time demands, it is generally necessary to wait until the end of ongoing activity before one is able to reflect on how the mistake happened and what could be changed to avoid a similar, future problem. In other cases, the professionals may notice the problem, such as internal bleeding in surgery, during the activity. If the bleeding problem is discovered hours later, it may be difficult for the professionals to remember what they were thinking at the time the actual problem occurred.

In an ideal learning environment based on simulations and re-created old cases, performers can wait until they are fully rested before confronting challenging situations. Based on our discussions of measurement of performance, students and advanced performers should be presented with cases just above their current level of ability. Ideally, the cases should also have known correct actions—where retrospective
analysis has reveals the best courses of appropriate action. The best training situations focus on activities of short duration with opportunities for immediate feedback, reflection, and corrections. Each completed trial should be followed by another similar brief task with feedback, until this type of task is completed with consistent success. At this point of mastery, the activity should be embedded in more complex contexts and alternated with other types of cases until the skill has been integrated in the performer’s repertoire. (For a more extended and detailed discussion on training in medical simulators and DP see the article by McGaghie.

CONCLUSION

Based on recent advances in the scientific analysis of reproducibly superior (expert) performance, we know that superior performance does not automatically develop from extensive experience, general education, and domain-related knowledge. Superior performance requires the acquisition of complex integrated systems of representations for the execution, monitoring, planning, and analyses of performance. Educators should therefore create training opportunities for DP, appropriate for a given individual at a given level of skill development. Performers may then make the necessary adjustments to improve specific aspects of performance to assure that attained changes will be successfully integrated into representative performance.

References

25. Ericsson KA, Williams AM. Capturing naturally-occurring superior performance in the laboratory.