Chapter 7
The Surgeon's Expertise

K. Anders Ericsson

7.1 Surgical Expertise: A Perspective from the Expert-Performance Approach

The emphasis on general education, problem-based training, and professional experience for the acquisition of skill, expertise, and professional achievement has varied during the history of training of professionals, such as engineers and medical doctors. As knowledge in the corresponding professional domain increased, it became clear that students had to attain a general education, such as a pre-medicine college education for doctors, before beginning their studies at their respective professional school. Following this primarily theoretical training, graduates were trained as apprentices and interns under the supervision of experienced practitioners for several years until they could earn the credentials to practice independently.

Traditional models of skill and expertise (Dreyfus and Dreyfus 1986; Fitts and Posner 1967) distinguish different phases of development of performance that are consistent with the distinction between general theoretical knowledge and professional skill. The first phase of the beginner, such as a medical student, involves reasoning from basic principles and then following instructions by teachers for applying step-by-step procedures. During this phase, gross errors occur and are noticed by the teacher, or even the student, and are corrected, and subsequently decrease in frequency. With increasing opportunities for performing similar tasks, the student becomes more able to generate better outcomes faster, more smoothly and with less effort. Some researchers of expertise (Dreyfus and Dreyfus 1986) consider individuals after extensive experience in the domain to become experts, who 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

K.A. Ericsson (✉)
Florida State University, Tallahassee, FL, USA
e-mail: ericsson@psy.fsu.edu

H. Fry and R. Kneebone (eds.), *Surgical Education: Theorising an Emerging Domain*, Advances in Medical Education 2, DOI 10.1007/978-94-007-1682-7_7,

we call expert" (Dreyfus and Dreyfus 1986, p. 21). In other more complex domains, such as telegraphy and chess, they argue that it may take over a decade to reach the highest levels (Simon and Chase 1973). The pioneering research on expertise by Herbert Simon and Bill Chase (1973) emphasized the improvements in performance associated with further experience in the domain and how increasingly complex patterns (or chunks) are acquired and stored in memory, providing the basis for pattern recognition to mediate rapid retrieval of appropriate actions from memory. Numerous studies in the late 1970s and early 1980s compared the performance of beginners with experts (Chi 2006; Feltovich et al. 2006). In these studies, it was common to identify experts by using peer-nomination procedures among highly experienced professionals (Elstein et al. 1978).

In the latter part of the 1980s, the conception was that accumulated knowledge of a domain, followed by an extended period of professional experience, would inevitably lead to expertise and superior performance, and peer nominations became increasingly criticized (see Chap. 1). Early studies of medical diagnosis were unable to establish superior accuracy of peer-nominated best general physicians compared to a group of undistinguished physicians (Elstein et al. 1978). Similar findings were subsequently obtained for clinical psychologists, where more advanced training and longer professional experience was unrelated to their success in treating patients' problems. Reviews show that there is a surprisingly weak relation between the length of professional experience and objective performance in a wide range of domains (Ericsson 2006a; Ericsson and Lehmann 1996). For example, the accuracy of heart sound diagnosis 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 as a function of years since graduation from training (Choudhry et al. 2005; Ericsson 2004; Ericsson et al. 2007).

It is important to note that in the majority of these domains, there is very little immediate feedback on the success or failure of a diagnosis. Many doctors never see the final diagnosis for a patient whom they try to diagnose, and if they do eventually see the diagnosis, their memory for their initial diagnostic process is too fragmentary to help assess what they overlooked or should have done. This situation is different in surgery, where mistakes, problems, and successful outcomes are often perceived during surgery, within hours of the completed surgery, or at least the next day, so accurate timely feedback is frequently available to help surgeons to learn and improve their skills. Consequently, as one of the exceptions from this general lack of learning from experience in professional domains, surgeons with more experience (larger number of completed surgical procedures of a given type) often have been found to have significantly superior outcomes for their patients (Ericsson 2004).

In response to this dissociation between superior performance and professional experience, Ericsson and Smith (1991) proposed that researchers should redirect research from studying socially recognized experts to studying reproducibly superior performance in a given domain.

7.2 The Scientific Study of Expert Performance and its Acquisition

The establishment of any science, including the study of expert performance, starts with accumulation of a body of reproducible empirical phenomena such as superior performance (Ericsson 2006a, b). Unless it is possible to reproduce such phenomena consistently, under standardized and experimental conditions, it will not be possible to analyze them with experimental methods.

The most successful efforts to demonstrate reproducibly superior performance under standardized conditions are found in sports. Athletic competitions in ancient Greece have a long history of attempting to design standardized situations that would allow fair competition between athletes. For example, they built straight and flat running tracks that were the same for all runners and devised methods to allow runners to start at the same time and then to cross the same finishing line—to make it easier to determine who got there first. More recently, competitions in music, dance, and chess, have been designed to evaluate the best performance with reliable and often objective methods for scoring. In all of these traditional domains, elite individuals reliably perform better than less accomplished individuals, when given technically difficult tasks.

There have also been efforts to measure performance in professional domains. In many of these domains, large numbers of professionals encounter and perform similar tasks on a daily basis. For example, professional investors on the stock market have essentially equal opportunities to purchase and sell stocks in companies; medical professionals, especially in emergency rooms of larger cities, treat patients with similar symptoms; and psychotherapists treat patients with similar reported problems. The number of encountered patients with the same type of challenging problems will be very small for a given professional, even if one were to aggregate their experience over a month or a year. There are, however, exceptions involving specialists, who are approached by people from a large area for help with particular procedures, such as specialized surgery. However, simply knowing that a surgeon has had better outcomes than other surgeons for some 50–100 patients (following surgical procedures that last around 4 h) makes it difficult to identify the nature and locus of the differences.

There is a similar problem in understanding what distinguishes chess masters from weaker players, when the games last several hours. In his pioneering research on chess expertise, de Groot (1946/1978) developed a methodology of selecting critical events. He extracted critical situations in games between chess masters and then set up a controlled laboratory situation where he could present the associated positions one at a time to an individual chess player (see Fig. 7.1). This method focuses on that part of a chess game where the masters' ability to select the best move in complex challenging situations is paramount to success. This excludes any differences in the beginning of a chess game, since these are typically simple and routine and are based on shared knowledge of openings. De Groot's method thus focuses on complex and challenging situations, where routines and prior experience will not lead to the best approach and actions.

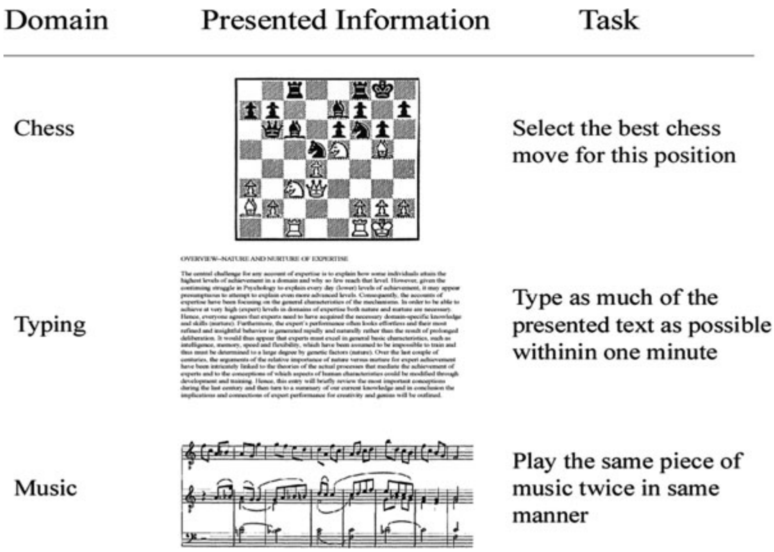

Fig. 7.1 Three examples of laboratory tasks that capture the consistently superior performance of domain experts in chess, typing, and music (From "Expertise," by Ericsson and Lehmann 1999)

For other types of tasks, the difficulty of selecting the correct actions may not be as important as the speed to complete typical tasks. Expertise in typing should be generalized to any type of material to be transcribed, so that one can present all typists with the same text material and ask them to type it accurately as fast as possible for a fixed time period. The final example given in Fig. 7.1 illustrates the skill of musical sight reading, where an accompanist would be presented with a sheet of music and asked to accompany a singer without having a chance to prepare the performance in advance. The ability to accurately play as many of the written notes as possible in a musically pleasing fashion is one characteristic that differentiates skilled accompanists from other pianists (Lehmann and Ericsson 1993). It is important to note that the speed of accurate sight reading is not an end in itself, but it demonstrates the musicians' ability or capacity for performance, which in turn allows them to musically express their pieces when they play at lower than maximal speed. Expert musicians are able to play a relatively simple piece twice in a row with a high degree of consistency, whereas novices are unable to reproduce the microstructure of their performance from trial to trial (Krampe and Ericsson 1996). Similarly, speed of surgeons' performance as well as their ability to reproduce their performance (displaying higher level of control) is likely to be correlated with surgical skill and surgical outcomes.

Over the last few decades, it has been possible to develop standardized test situations, where performance on these representative situations can be assessed in around an hour, yet are highly correlated to real-world performance. Examples include tournament performance in chess, golf, and Scrabble, performance in music competitions, and medical diagnosis (Ericsson 2006a, b). 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. In many domains, there is a rank ordering of difficulty for mastery of different tasks. For example, in diving competitions each dive is assigned a difficulty score, while in music, pieces are rated by the number of years of study recommended before attempted mastery. In addition, gymnastics and martial arts have a clear progression of levels defined by mastery of increasingly difficult tasks, and a similar progression of mastery is found in mathematics and many of the sciences.

7.2.1 Applying the Expert-Performance Approach to Surgery

The central question for the expert-performance approach when applied to surgery concerns if, and under what circumstances, it is possible to identify surgeons who have consistently superior outcomes for their patients—although it may sometimes be difficult to separate the effects of the individual surgeon from the contributions of their team members and the influence of post-surgery care. Superior data on patient outcomes is the most compelling for procedures with stable individual differences in objective outcomes, such as mortality and morbidity. Recently, Vickers et al. (2007) reported large differences in mortality as a function of the number of procedures of that same type previously completed by the surgeon—patients whose surgeon had performed less than ten procedures were almost twice as likely to have a recurrence of the cancer as patients whose surgeons had performed more than 250 procedures. The biochemical recurrence following this particular procedure (removal of the prostate) is claimed by Vickers et al. (2007, p. 1171) to be a particularly good measure on surgical performance "because adjuvant therapy is not commonly given for prostate cancer and recurrence is not substantially affected by other aspects of postoperative care." With statistical control for severity of the patients' cancer condition, it is possible to determine that these differences in outcome are due to acquired skill. In fact, Vickers et al. (2007) compared the same surgeons' outcomes for their first 10 procedures and after 250 procedures and observed the same significant improvement. In a subsequent study, Vickers et al. (2008) discovered that when cancers that were confined to an organ, the recurrence of cancer was monotonically reduced with increased surgeons' experience for the first 1,500–2,000 procedures to a point at which recurrence of cancer was essentially eliminated. In a different analysis of outcomes of laparoscopic procedures, Vickers et al. (2009) found that decreases in recurrence of cancer were seen for the first 1,000 procedures, whereas with open surgery a stable plateau was attained after 250 procedures. More generally, individual differences in outcomes for practicing board-certified surgeons are often found for the more complex procedures with the highest mortality rates. For example, Prystowsky (2005) found no differences for mortality as a function of number of procedures for simple cases of alimentary tract surgery (ATS), but only for complex ATSs. He reviews seven other studies demonstrating a higher mortality or complication rate for the first 15–50 cases of complex procedures by certified surgeons. Recent reviews of this research demonstrate that surgical performance minimizing mortality and morbidity requires procedure-extended specific experience and training, and that general surgeons require considerable training in laparoscopic techniques to overcome the steep learning curve (Kumar and Gill 2006; Prystowsky 2005; Vickers et al. 2009).

Several successful approaches can be adopted for identifying superior (expert) performance. For example, it is possible to study surgical teams which differ in their associated risk-adjusted mortality rates, then have all teams adopt the procedures of best-performing teams. This adoption of expert methods with frequent monitoring of units' outcomes has led to significant improvements in system-wide mortality outcomes (Nugent 2005).

It is much more difficult to analyze the infrequent instances leading to mortality that measure and give immediate feedback on the performance of individual surgeons. For example, recurrence of some types of cancers will happen up to 1–5 years after the operation, thus precluding immediate feedback after surgery. It is, however, possible to monitor the detailed processes of a given operation by video recording or even having a surgeon "think aloud" while performing the operation. Frequently, the surgeons discover mistakes themselves or other staff notice problems during or after surgery. In addition, it is possible to gain additional feedback about the surgical outcome using special tests. For example, following radical prostatectomy, Atug et al. (2006) analyzed the tissue removed during surgery to assess whether the edges included cancerous tissue or whether the cancerous tissue was completely contained within the removed tissue. They found that with increased surgical experience, the number of collected samples with cancerous tissue decreased by a factor of 4. Similarly, Bacha et al. (2008) describe how the outcomes of congenital heart surgery can be evaluated almost immediately after the surgery by post-procedure echocardiographic testing that assesses the repaired heart's function. By developing similar auxiliary tests, conducted for the primary purpose of giving feedback on outcomes within hours or days after the completed surgery, the continued learning and improvement of surgeons would be facilitated.

The most promising approach to studying individual differences in surgical performance and its relation to surgical experience involve videotaping actual surgeries operations, followed by blind assessment of the surgeons' performance. There are few studies that have compared live operations performed by highly experienced surgeons who have completed training (internship and residency). In a pioneering study, Sarker et al. (2006) compared video tapes from laparoscopic cholecystectomies by four less-experienced registrar/resident surgeons with those by five experienced consultant/attending surgeons, and found significantly faster and higher ratings for general and specific performance for the more experienced experts. In a more recent study, Murphy et al. (2008) completed a task analysis to identify serious errors, such as organ perforations and tissue tearings. They then assessed the error rates for novices (interns and junior registrars) and experts (senior specialist registrars and consultants) and found significantly lower rates for the more expert group. These findings should permit better targeted training on the development of control of the movements with one's instrument, and inform the design of training activities outside the operating theater that would permit repeated attempts with immediate feedback to develop that control.

There has not been any research requiring surgeons to "think aloud" during a particular surgery or to give a retrospective report immediately after the surgery in order to link individual differences in thinking to superior surgical outcomes. One of the few "think-aloud" studies was conducted by Abernathy and Hamm (1994), who asked a master surgeon to "think aloud" about how to treat sick patients in different scenarios. The focus, however, was not on the planning or execution of surgery but on general diagnosis and treatment (see Ericsson 2004, 2007, for discussions of diagnosis of patient vignettes and simulated patients). The methodology that most closely resembles an application of the expert-performance approach to surgery is illustrated in a recent study by Sarker et al. (2009). They relied on a library of videotapes of over 100 operations to identify situations requiring a decision, where two highly experience laparoscopic surgeons agreed on the correct decision. These situations were presented to experts and intermediate surgeons, and the expert group made significantly better surgical decisions.

Another interesting approach involves inviting experienced surgeons to perform simple tasks used for training surgical students, to provide reference points for mastery of six laparoscopic tasks. An established simulator is the Minimally Invasive Surgical Trainer-Virtual Reality (MIST-VR), a device simulating laparoscopic surgery by using realistic instruments to perform geometrical tasks on computed visual images. Van Sickle et al. (2007) tested over 40 experienced surgeons (who had an average of over 1,000 laparoscopic procedures) on the MIST-VR after a single training trial for each laparoscopic task. Completion time and errors were recorded for each task, but showed no significant correlations with years of laparoscopic experience or number of previous laparoscopic procedures. The lack of correlation with experience was not due to any ceiling effects, as the surgeons' previous experience with the particular simulator MIST-VR was associated with superior performance.

In sum, it is possible to identify individuals whose performance is consistently superior to that of other individuals in the same domain. In particular, in the domain of surgery, the number of times a particular surgeon has completed a given procedure is a potent predictor of surgical outcome for complex and challenging procedures. In the following sections, I will discuss how we can learn from the expert performers and their developmental path to superior performance.

7.3 The Acquisition of Superior Reproducible (Expert) Performance

In most of the traditional domains of expertise, such as chess, sports, and music, it has been possible to describe the time course of development that generalizes across different domains of expertise (Ericsson 2006a, b; Ericsson and Lehmann 1996).

Unlike surgery, most international level performers in the traditional domains start their training and practice as children, often around 7–8 years of age but sometimes as early as 3–4 years. Even though surgical training starts two decades later, similar characteristics are observed.

When we are using the same objective standards to measure performance, such as chess-ratings, time to complete running events, and risk-adjusted mortality rates for surgical outcomes, there is no evidence for abrupt increases in performance, and learning curves show gradual smooth improvement. While analyses of performance in sport and chess tend to show that the age at which experts typically reach their peak career is in the 30s and 40s, analyses of surgical performance show a pattern similar to music (Krampe and Ericsson 1996) with no reliable decline for active professionals (Waljee et al. 2006). Finally, in other domains (including sports, sciences, and arts), researchers have found that all performers, even the most "talented," need around 10 years of intense involvement before they reach an international level (Ericsson et al. 1993; Simon and Chase 1973), and most elite individuals take considerably longer. In sports and even music there are regional, national, and international competitions to assess when someone is able to win at a given level of competition. In surgery, by contrast, there are no established competitions and associated measures of performance, so the most relevant evidence would be found in data on how long it takes to reach the lowest mortality rates for surgeries with high base rates for mortality, such as advanced forms of cancer. Consistent with the need for up to a decade of required engagement in domain-related activities, the best surgical performance is attained after 1,500–2,000 operations (Vickers et al. 2008) or over 7 years of experience in the emergency department (McKenney et al. 2009). Surgical skill is not a stable entity and is constantly in flux, with technical innovations and the new surgical techniques being constantly introduced. Today's surgeons need to acquire an ability to keep learning new methods and techniques during their professional careers.

7.3.1 The Acquisition of Surgical Expertise

The primary challenge for surgical training has been to handle the steep initial learning curve. Traditionally, surgical trainees were allowed to perform increasingly complex aspects of a surgical procedure under the direct supervision of a trained surgeon. The experienced surgeon would step in and handle any problems or difficult parts of the surgery until the trainee had acquired a proficient performance. Research has shown that trainees performing surgeries under the supervision of surgeons experienced in that procedure have mortality rates similar to surgeries performed by experienced surgeons (Stoica et al. 2008). This is the general method of apprenticeship training that has dominated surgical education until quite recently, when the weekly hours of surgical service by residents have been markedly reduced.

More recently introduced surgical techniques that differ radically from the traditional open operations, such as laparoscopic and robotic procedures, have led to a challenge. As reviewed earlier, when certified practising surgeons initially perform procedures for which they lack experience, there is a higher incidence of mortality and other postoperative complications, which is reduced after 30–100 completed operations. Consequently, surgeons need supervised training before they are able to independently complete operations at a proficient level. Given the diverse and large number of procedures and the need for lengthy supervised training for each procedure prior to the acquired proficiency, there has evolved a great interest in training in simulators that would allow skill acquisition outside the operating theater. Some recent research has suggested that expert performers in many domains, such as music, acting, and sports, have developed training methods that do not require the expert to be in front of an audience but allows the training to occur during rehearsal, instruction by a coach or teacher, and solitary practice.

In a review of research on skill acquisition in a wide range of domains, my colleagues and I (Ericsson et al. 1993) identified a set of conditions where practice had consistently led to increased performance. Individuals who were given a task with a well-defined goal were motivated to improve, were provided with feedback, and had ample opportunities for gradual refinements of their performance with repetition of the same or similar tasks, saw their performance improve significantly. Deliberate efforts to improve performance beyond its current level often require identification of those aspects of performance requiring improvement and finding better methods to perform the tasks. These activities demand full concentration, limiting daily duration (Ericsson 2006a, b).

Informal observation of learning of everyday skills, such as tennis, golf, typing on a computer, and using a mobile phone, shows that people initially figure out what to do slowly, but with more practice opportunities are able to reach a sufficient level in the target activity (such as sending e-mails or text messages or returning a tennis shot). Many people spontaneously adopt the most available strategies, such as hunt-and-peck in typing or idiosyncratic movement patterns in tennis. 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 (Dreyfus and Dreyfus 1986; Fitts and Posner 1967) (Fig. 7.2, lower arm). When performance has reached a level of automaticity, additional experience will not improve the structure or accuracy of action selection, and consequently the amount of accumulated experience will not be related to attained level of performance.

In direct contrast, teachers help aspiring experts to adopt the best training methods and the appropriate fundamentals, gradually developing into expert performance without long periods of relearning poor fundamentals. By helping their students target weaker aspects of their performance with deliberate practice, improvements toward the expert performance is faster than if practice involves unstructured experience without explicit goals. In music and sports, most of the training is not completed during informal games, competitions and public performances, but takes place on the practice field or the practice room.

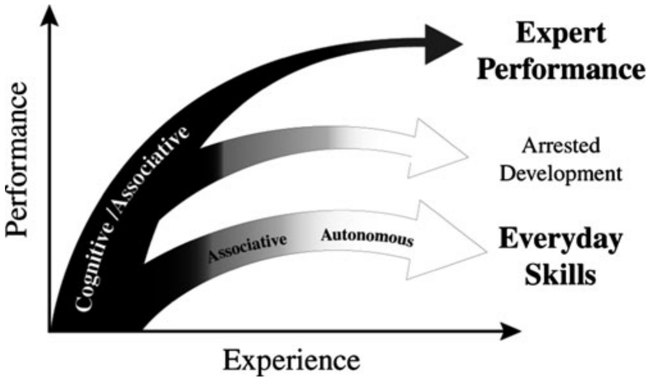

Fig. 7.2 An illustration of the qualitative difference between the course of improvement of expert performance and of everyday activities. The goal of everyday activities is to reach as rapidly as possible a satisfactory level that is stable and "autonomous." After individuals pass through the "cognitive" and "associative" phases they can generated their performance virtually automatically with a minimal amount of effort (see the *gray*/*white* plateau at the bottom of the graph). In contrast, expert performers counteract automaticity by developing increasingly complex mental representations to attain higher levels of control of their performance and will therefore remain within the "cognitive" and "associative" phases. Some experts will, at some point in their career, give up their commitment to seeking excellence and thus terminate regular engagement in deliberate practice to further improve performance, which results in premature automation of their performance (Adapted from "The scientific study of expert levels of performance: General implications for optimal learning and creativity" by Ericsson (1998). Copyright 1998 by European Council for High Ability)

7.3.2 Simulation

Simulation offers obvious potential here (see Chaps. 3 and 8). A recent review (Issenberg et al. 2005) has shown that the improvements in performance due to simulator training are primarily seen for training involving explicit performance goals with opportunities for feedback and repetitions until mastery (training activities satisfying the characteristics of deliberate practice). In a subsequent review, McGaghie et al. (2006) showed that the amount of deliberate practice showed a dose–response relation to performance. There has been a virtual explosion of research on the design of effective simulator training for surgical performance, especially for laparoscopic and robotic procedures. Recent comprehensive reviews (McGaghie 2008; Tsuda et al. 2009) note that education in simulators embraces the best methods, such as "distributed, structured, and deliberate practice with the appropriate mechanisms for feedback" (Tsuda et al. 2009, p. 336), with objective training goals based on experienced surgeons' performance in the simulator (c.f. Van Sickle et al.'s (2007) testing of expert surgeons in the simulators).

It is interesting to note that medical educators are adopting the characteristics of training (deliberate practice) originally observed among expert musicians (Ericsson et al. 1993). In music, the teacher identifies an aspect of the students' music performance that needs to be improved and then recommends particular training methods and techniques, where the targeted aspect can be gradually refined through repetition and refinement in response to feedback. The music student engages in the assigned practice activities until the goal is attained or the student is no longer able to engage in full concentration on their skill acquisition. The constraint on maintaining full concentration leads the students to limit training session to an hour and no more than 4–5 h of deliberate training each day.

There are now several impressive demonstrations of improvement in performance of surgical simulators as well as the transfer of simulator training to the operating room (Seymour 2007). For example, Ahlberg et al. (2007) showed a reliable decrease in errors for simulator-trained residents during their first ten laparoscopic cholecystectomies in the OR. In surgery and many other domains of traditional expertise, such as music, ballet, and sports, beginners need the help of teachers to identify appropriate aspects that are especially amenable to improvement. The teachers also are critical in helping students identify appropriate training techniques that lead to the desired goals within hours of training. Teachers are particularly important in evaluating and monitoring performance until the students eventually develop the skills to be able to monitor their own performance and become their own teacher.

In numerous domains, such as chess, music, and sports, aspiring experts acquire memory representations that allow them to rapidly encode situations and to evaluate and plan their future actions (Ericsson 2006a). Similarly, in surgery residents develop mental representations to support their ongoing evaluation and planning. For example, Bann et al. (2005) found a high correlation between residents' ability to detect errors in models and their ability to complete the same procedure in the operating room. In additional support for such representations, Wiegmann et al. (2007) found that residents discover most of their mistakes during surgery, but that interruptions of the operation by external factors, such as telephone calls, lead to increased probability of errors. Way et al. (2003) give examples of the challenges in identifying the anatomical structures during laparoscopic surgery and the associated skills allowing experienced surgeons to reduce the risks of injuring adjacent tissues, ducts, and vessels.

Once residents have completed their training and achieved their certification, they should remain motivated to continue improving and maintaining their skills. Unfortunately, some surgeons may develop automaticity during their practice (Fig. 7.2, middle arm). Consistent with such a development, Bann et al. (2005) argued that "senior surgeons are more prone to slips and lapses" (p. 414). Consequently, the key challenge for aspiring elite performers in any domain of expertise is to avoid the arrested development associated with automaticity. Individuals striving for excellence need actively to counteract tendencies toward automaticity (Fig. 7.2, upper arm). They do that by setting new and higher standards for their performance, requiring them to increase their speed, accuracy, and control over their action generation. For example, surgeons can assess their surgical margins (in cancer surgery), try to reduce redundant movements, and increase the safety and control of their movements by retrospective analysis of video tapes.

Experts deliberately construct and/or seek out training situations in which they can stretch themselves to attain desired goals that exceed their current level of reliable performance. They acquire and refine mechanisms that permit increased control and allow them to monitor performance in representative situations from the domain of expertise, so they can identify errors as well as improvable aspects (Ericsson 2006a, b).

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 during the move selection and also recall the locations of all the pieces on the chess board virtually perfectly. Experts' superior incidental memory for relevant information for representative tasks has been demonstrated in a large number of domains, such as sports, music, ballet, and medicine (Ericsson and Kintsch 1995; Ericsson et al. 2000). When expert performers are given appropriately challenging tasks then they have to think, image, and reason. The most direct evidence for this type of thinking comes from asking the expert to "think aloud" during the procedure or to give a retrospective report on their thoughts immediately following the procedure (Ericsson 2006a).

7.4 Concluding Remarks

In this chapter, I have suggested parallels between findings on expert performance in surgery and those in other domains. From laboratory analyses of experts' superior performance in traditional domains, scientists have consistently found evidence for the acquired mediating mechanisms discussed above: very complex skills, highly refined representations, and extreme physiological adaptations to physical domains. In this chapter, I have tried to show how the acquisition of superior performance in surgery is closely related to the extent of engagement in practice with feedback during medical training and residency. I have also speculated that after the end of organized medical training, continued access to conditions for deliberate practice as well as feedback on daily medical practice might allow surgeons to keep improving their performance.

The complex integrated structure of expert performance raises many issues about how these structures can be gradually acquired and perfected over time. Medical students need to acquire representations that can support their planning, reasoning, and evaluation of the actual and intended performance to be able to make more appropriate adjustments to their complex skills (see Chaps. 3, 8, 10 and 12). This advantage becomes absolutely essential at higher levels of achievement. Given that deliberate practice involves mastering tasks that students could not initially attain, or only attain imperfectly or unreliably, successful students seem to acquire the ability to think, plan, and reason; this ability is further refined to allow them to solve problems and learn distinctions and consequences through planning and analysis.

In sum, I believe that the study of expert performance in surgery and other areas of medicine will provide unique insights for how to apply the expert performance framework to the study of many types of professional expertise. I anticipate that future research will show that the promising application of the expert-performance approach to medicine will advance our understanding of the development of professional expertise, and thus will yield measurable improvements in the performance of experts in many professional domains in our society.

References

Abernathy, C. M., & Hamm, R. M. (1994). *Surgical scripts: Master surgeons think aloud about 43 common surgical problems*. Philadelphia: Hanley & Belfus.

Ahlberg, G., Enochsson, L., Gallagher, A. G., Hedman, L., Hogman, C., McClusk, D. A., Ramel, S., Smith, C. D., & Arvidsson, D. (2007). Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. *American Journal of Surgery, 193*, 797–804.

Atug, F., Castle, E. P., Srivastav, S. K., Burgess, S. V., Thomas, R., & Davis, R. (2006). Positive surgical margins in robotic-assisted radical prostatectomy: Impact of learning curve on oncologic outcomes. *European Urology, 49*, 866–871.

Bacha, E. A., Larrazabal, L. A., Pigula, F. A., Gauvreau, K., Jenkins, K. J., Colan, S. D., Fynn-Thompson, F., Mayer, J. E., Jr., & del Nido, P. J. (2008). Measurement of technical performance in surgery for congenital heart disease: The stage I Norwood procedure. *The Journal of Thoracic and Cardiovascular Surgery, 136*, 993–997.

Bann, S., Khan, M., Datta, V., & Darzi, A. (2005). Surgical skill is predicted by the ability to detect errors. *American Journal of Surgery, 189*, 412–415.

Chi, M. T. H. (2006). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 21–38). Cambridge: Cambridge University Press.

Choudhry, N. K., Fletcher, R. H., & Soumerai, S. B. (2005). Systematic review: The relationship between clinical experience and quality of health care. *Annals of Internal Medicine, 142*, 260–273.

de Groot, A. (1978). *Thought and choice in chess*. The Hague: Mouton (Original work published 1946).

Dreyfus, H., & Dreyfus, S. (1986). *Mind over machine: The power of human intuition and expertise in the era of the computer*. New York: Free Press.

Elstein, A. S., Shulman, L. S., & Sprafka, S. A. (1978). *Medical problem solving: An analysis of clinical reasoning*. Cambridge: Harvard University Press.

Ericsson, K. A. (1998). The scientific study of expert levels of performance: General implications for optimal learning and creativity. *High Ability Studies, 9(1)*, 75–100.

Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine, 79*, S70–S81.

Ericsson, K. A. (2006a). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 685–706). Cambridge: Cambridge University Press.

Ericsson, K. A. (2006b). Protocol analysis and expert thought: Concurrent verbalizations of thinking during experts' performance on representative task. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 223–242). Cambridge: Cambridge University Press.

Ericsson, K. A. (2007). An expert-performance perspective on medical expertise: Study superior clinical performance rather than experienced clinicians! *Medical Education, 41*, 1124–1130.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211–245.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence on maximal adaptations on task constraints. *Annual Review of Psychology, 47*, 273–305.

Ericsson, K. A., & Lehmann, A. C. (1999). *Encyclopedia of creativity*. San Diego: Academic Press.

Ericsson, K. A., & Smith, J. (1991). Prospects and limits in the empirical study of expertise: An introduction. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 1–38). Cambridge: Cambridge University Press.

Ericsson, K. A., Krampe, R. Th., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363–406.

Ericsson, K. A., Patel, V. L., & Kintsch, W. (2000). How experts' adaptations to representative task demands account for the expertise effect in memory recall: Comment on Vicente and Wang (1998). *Psychological Review, 107*, 578–592.

Ericsson, K. A., Whyte, J., & Ward, P. (2007). Expert performance in nursing: Reviewing research on expertise in nursing within the framework of the expert-performance approach. *Advances in Nursing Science, 30*, E58–E71.

Feltovich, P., Prietula, M. J., & Ericsson, K. A. (2006). Studies of expertise from psychological perspectives. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 41–67). Cambridge: Cambridge University Press.

Fitts, P., & Posner, M. I. (1967). *Human performance*. Belmont: Brooks/Cole.

Issenberg, S. B., McGaghie, W. C., Petrusa, E. R., & Scalese, R. J. (2005). Features and uses of high-fidelity medical simulations that lead to effective learning: A BEME systematic review. *Medical Teacher, 24*, 10–28.

Krampe, R. Th., & Ericsson, K. A. (1996). Maintaining excellence: Deliberate practice and elite performance in young and older pianists. *Journal of Experimental Psychology. General, 125*, 331–359.

Kumar, U., & Gill, I. S. (2006). Learning curve in human laparoscopic surgery. *Current Urology Reports, 7*, 120–124.

Lehmann, A. C., & Ericsson, K. A. (1993). Sight-reading ability of expert pianists in the context of piano accompanying. *Psychomusicology, 12*(2), 182–195.

McGaghie, W. C. (2008). Research opportunities in simulation-based medical education using deliberate practice. *Academic Emergency Medicine, 15*, 995–1001.

McGaghie, W. C., Issenberg, S. B., Petrusa, E. R., & Scalese, R. J. (2006). Effect of practice on standardised learning outcomes in simulation-based medical education. *Medical Education, 40*, 792–797.

McKenney, M. G., Livingstone, A. S., Schulman, C., Stahl, K., Lineen, E., Namias, N., et al. (2009). Trauma surgeon mortality rates correlate with surgeon time at institution. *Journal of the American College of Surgeons, 208*, 750–754.

Murphy, D., Doherty, G., & Luz, S. (2008). Differentiating between novices and expert surgeons based on errors derived from task analysis. *ECCE '08 Proceedings of the 15th European conference on Cognitive ergonomics: the ergonomics of cool interaction*. New York, NY, USA: Association for Computing Machinery.

Nugent, W. C. (2005). Building and supporting sustainable improvement in cardiac surgery: The Northern New England experience. *Seminars in Cardiothoracic and Vascular Anesthesia, 9*, 115–116.

Prystowsky, J. B. (2005). Are young surgeons competent to perform alimentary tract surgery? *Archives of Surgery, 140*, 495–502.

Sarker, S. K., Chang, A., Vincent, C., & Darzi, A. W. (2006). Development of assessing generic and specific technical skills in laparoscopic surgery. *American Journal of Surgery, 191*, 238–244.

Sarker, S. K., Rehman, S., Ladwa, M., Chang, A., & Vincent, C. (2009). A decision-making learning and assessment tool in laparoscopic cholecystectomy. *Surgical Endoscopy, 23*, 197–203.

Seymour, N. E. (2007). VR to OR: A review of the evidence that virtual reality simulation improves operating room performance. *World of Surgery, 32*, 182–188.

Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist, 61*, 394–403.

Stoica, S. C., Kalavrouziotis, D., Martin, B.-J., Buth, K. J., Hirsch, G. M., Sullivan, J. A., et al. (2008). Long-term results of heart operations performed by surgeons-in-training. *Circulation, 118*, S1–S6.

Tsuda, S., Scott, D., Doyle, J., & Jones, D. B. (2009). Surgical skills training and simulation. *Current Problems in Surgery, 46*, 271–370.

Van Sickle, K. R., Ritter, E. M., McClusky, D. A., III, Lederman, A., Baghai, M., Gallagher, A. G., & Smith, C. D. (2007). Attempted establishment of proficiency levels for laparoscopic performance on a national scale using simulation: The results from the 2004 SAGES Minimally Invasive Surgical Trainer-Virtual Reality (MIST-VR) learning center study. *Surgical Endoscopy, 21*, 5–10.

Vickers, A. J., Bianco, F. J., Serio, A. M., Eastham, J. A., Schrag, D., Klein, E. A., et al. (2007). The surgical learning curve for prostate cancer control after radical Prostatectomy. *Journal of the National Cancer Institute, 99*, 1171–1177.

Vickers, A. J., Bianco, F. J., Gonen, M., Cronin, A. M., Eastham, J. A., Schrag, D., et al. (2008). Effects of pathologic state on the learning curve for radical prostatectomy: Evidence that recurrence in organ-confined cancer is large related to inadequate surgical technique. *European Urology, 53*, 960–966.

Vickers, A. J., Hruza, M., Koenig, P., Martinez-Pinero, L., Janetschenk, G., & Guillonneau, B. (2009). The surgical learning curve for laparoscopic radical prostatectomy: A retrospective cohort study. *The Lancet Oncology, 10*, 475–480.

Waljee, J. F., Greenfield, L. J., Dimick, J. B., & Birkmeyer, J. D. (2006). Surgeon age and operative mortality in the United States. *Annals of Surgery, 244*, 353–362.

Way, L. W., Stewart, L., Gantert, W., Liu, K., Lee, C. M., Whang, K., & Hunter, J. G. (2003). Causes and prevention of laparoscopic bile duct injuries. Analysis of 252 cases from a human factors and cognitive psychology perspective. *Annals of Surgery, 237*, 460–469.

Wiegmann, D. A., ElBardissi, A. W., Dearani, J. A., Daly, R. C., & Sundt, T. M. (2007). Disruptions in surgical flow and their relationship to surgical errors: An exploratory investigation. *Surgery, 142*, 658–665.