Part II

OVERVIEW OF APPROACHES TO THE STUDY OF EXPERTISE – BRIEF HISTORICAL ACCOUNTS OF THEORIES AND METHODS
CHAPTER 4

Studies of Expertise from Psychological Perspectives

Paul J. Feltovich, Michael J. Prietula, & K. Anders Ericsson

Introduction

The study of expertise has a very long history that has been discussed in several other chapters in this handbook (Ericsson, Chapter 1; Amirault & Branson, Chapter 5). This chapter focuses on the influential developments within cognitive science and cognitive psychology that have occurred over the last three decades. Our chapter consists of two parts. In the first part we briefly review what we consider the major developments in cognitive science and cognitive psychology that led to the new field of expertise studies. In the second part we attempt to characterize some of the emerging insights about mechanisms and aspects of expertise that generalize across domains, and we explore the original theoretical accounts, along with more recent ones.

The Development of Expertise Studies

In this handbook there are several pioneering research traditions represented that were brought together to allow laboratory studies of expertise, along with the development of formal models that can reproduce the performance of the experts. One early stream was the study of thinking using protocol analysis, where participants were instructed to “think aloud” while solving everyday life problems (Duncker, 1945), and experts were asked to think aloud while selecting moves for chess positions (de Groot, 1946/1965; Ericsson, Chapter 13). Another stream developed out of the research on judgment and decision making, where researchers compared the judgments of experts to those of statistical models (Meehl, 1954; Yates & Tschirhart, Chapter 24). The most important stream was one inspired by describing human performance with computational methods, in particular, methods implemented as programs on the computer, such as Miller, Galanter, and Pribram (1960), Reitman (1965), and Newell and Simon, (1972).

In this chapter we emphasize a period of research roughly from the mid-1950s into the 1970s, when empirical experimental studies of thinking in the laboratory were combined
with theoretical models of human thought processes that could reproduce the observable performance. Even though there was important earlier work on expertise, this was the period when a number of forces came together to provide enough traction for the field to “take off.” There were three main sources to this impetus: artificial intelligence, psychology, and education. We will survey these briefly.

Early computer models developed by Herbert Simon and Allen Newell demonstrated that it is relatively easy for computational devices to do some things worthy of being considered “intelligent.” This breakthrough at Carnegie-Mellon was based on the confluence of two key realizations that emerged from the intellectual milieu that was developing between Carnegie and Rand at the time (Prietula & Augier, 2005). First, they (Al Newell, Cliff Shaw, and Herb Simon) envisioned that computers could be used to process “symbols and symbol structures.” To explore this, they necessarily developed what was to become the first list-processing computer language, IPL, which afforded them the ability to create arbitrarily complex list structures and manipulate them recursively. Second, they incorporated the concept of “levels of abstraction” in articulating their theories and, consequently, their programs. These allowed them to address two critical technical problems: the “specification problem,” in which the components and processes of the target system are sufficiently specified to capture the characteristics of interest, and the “realization problem,” in which the specification can be implemented in an actual physical system to enable synthesis (Newell & Simon, 1956). The seeds of viewing humans and machines as complex information-processing systems had been sown.

During these early years, the first artificial intelligence program, called the Logic Theorist (Newell & Simon, 1956), was written. The Logic Theorist (LT) was coded in IPL. Significantly, it was able to prove theorems in the predicate calculus in a manner that mimics human adults (Newell & Simon, 1972). Of particular relevance to expertise, LT was able to create some novel proofs. The heuristics from LT were later generalized into a model that could solve problems in many different domains, the General Problem Solver (Ernst & Newell, 1969). There were also other computer models that were built, not as simulations of human problem solving, but based on effective computation designed to represent artificial methods for producing intelligent action. For example, Samuel’s (1959) checker-playing program was able to challenge and beat excellent human checker players. These early, along with subsequent, successes spawned some themes regarding expertise pertinent to the present chapter.

First, the idea that computation could support intelligent behavior reinforced the growing idea that computers and their programs could stand as formal models of human cognition. This grew into a pervasive stance toward human and machine cognition, the “information processing” model that is still widely held. Cognitive psychology and computer science merged into a very close collaboration (along with linguistics and a few other fields) that was later named Cognitive Science. These computational models and theories provided at least alternatives to the “behaviorist” (stimulus-response, no internal mental mechanisms) approaches that had dominated psychology for the prior half a century (more on this in our treatment of psychology and expertise below). Newell and Simon, two pioneers of the information-processing viewpoint, asserted this forcefully:

As far as the great debates about the empty organism, behaviorism, intervening variables, and hypothetical constructs are concerned, we take these simply as a phase in the development of psychology. Our theory posits internal mechanisms of great extent and complexity, and endeavors to make contact between them and the visible evidences of problem solving. That is all there is to it. (Newell & Simon, 1972, pp. 9–10)

As we will address in our treatment of psychological influences, it is quite difficult to imagine what a field of studying expertise
could have looked like if behaviorism had continued to hold sway.

The second theme has to do with alternative basic approaches to achieving intelligence in a computational device, what have been termed “weak and strong methods” (Newell, 1973). The earliest successful AI programs utilized weak reasoning and problem-solving methods that were drawing on descriptions of human thought processes. Indeed, at one point Newell termed artificial intelligence the “science of weak methods,” at least as one characterization of AI (Newell, 1973, page 9). Weak methods are highly portable, generalizable methods that do not depend on the particular content of the domain of problem solving but, in being so, are less capable of finding solutions. Examples are “generate and test” (produce and apply all possible known next steps, and see if any of them yields success) and “means-ends analysis” (represent the goal state, what you are trying to achieve; represent where your progress has brought you right now; and try to find some currently available computational operator that can decrease some aspect of the distance between these. Repeat until done. Strong methods are more heavily dependent on rich knowledge of the problem-solving area and an understanding of what kinds of operations are likely to be successful in encountered situations. They are domain specialists, not generalists.

When early AI was being applied in relatively simple and well-structured areas, such as elementary games like checkers, weak methods fared fairly well. As the field developed and researchers started to address richer, complex, and knowledge-laden task environments, such as medicine (Pauker, Gorry, Kassirer, & Schwartz, 1976; Shortliffe, 1976) and chemical spectral analysis (Buchanan & Feigenbaum, 1978), the need for ever-stronger methods became clear. Portability across task domains had to be sacrificed in favor of capability, but narrowly restricted capability. The highly successful “expert systems” industry that eventually developed (Buchanan, Davis, & Feigenbaum, Chapter 6) is in large part testimony to the efficacy of strong methods. As related to this chapter, this is important because a similar progression unfolded in other kinds of investigations of expertise, including those in psychology (see later sections in this chapter on “Expertise Is Limited in Its Scope and Elite Performance Does Not Transfer”; and “Knowledge and Content Matter Are Important to Expertise”).

Behaviorism was the school of psychology that eschewed resorting to unobservable mental constructs, structural or process, of any kind. Only the observable environment (the stimulus) and an organism’s overt reaction (the response) were considered the legitimate purview of a psychological science. Behaviorism had dominated psychology for much of the first half of the twentieth century. During the reign of behaviorism, considerable success was obtained in analyzing complex skills in terms of acquired habits, that is, as a large collection of stimulus-response pairs in the form of learned reactions associated to specific situations. The principle difficulties of this approach were associated with explaining the acquisition of abstract rules, creative use of language, general mental capacities, and logical reasoning in unfamiliar domains. It was around the middle of the century that this hold on the field began to loosen. There was both a push side and a pull side to this development.

On the push side, as we have noted, stimulus-response models were facing great difficulty in trying to account for complex human processes such as language, reasoning, and abstractions that were independently coming under increasing investigation. In this respect, the work of the linguist Chomsky (1957) was critical. The findings and theorizing out of linguistics were affecting psychology, in exposing what seemed to be significant inadequacies in accounting for complex psychological processes. A notable volume (Jakobovits & Miron, 1967), not surprisingly focusing on language, brought the camps head to head in their explanatory systems for complex human activity. The Herculean effort by Osgood (1963), reprinted in that volume, to save S-R theory
in the face of discoveries about language, just served in its cumbersomeness to prove the inadequacies of S-R theories to account for language.

On the pull side, theories, mechanisms, and constructs were arising that showed promise for providing an infrastructure to support a new kind of psychology. These included the development of the information-processing viewpoint in psychology, along with the platform to support it, the computer. Electrical engineer Newell and economist/philosopher Simon believed that what they were doing was psychology (see earlier quote)! In fact, they predicted in 1958 that “within ten years most theories in psychology will take the form of computer programs, or of qualitative statements about the characteristics of computer programs” and discussed the nature of heuristic search and ill-structured problems (Simon & Newell, 1958, p. 7). In his landmark volume titled “Cognitive Psychology,” Ulric Neisser (1967) engaged information-processing language and the computer metaphor as advances that helped enable the creation of a cognitive psychology, and he acknowledged the contributions of Newell, Shaw, and Simon in this regard (Neisser, 1967, pp. 8–9).

Additionally, and often not independently, researchers were progressively encroaching the realm of the mental, studying such things as planning (Miller, Galanter, & Pribram, 1960), thinking (Bartlett, 1958; Bruner, Goodnow, & Austin, 1956), and mental structures and their functioning (Bartlett, 1932; Miller, 1956). Not surprisingly, groundbreaking progress in this regard came from the information-processing camp in their studies of problem solving (Newell & Simon, 1972), especially in their studies (following de Groot, 1946, 1965) of expertise in chess (Chase & Simon, 1973a, 1973b; See also Gobet & Charness, Chapter 30). The clear, surprising, and even enchanting findings (two people looking at the very same “stimulus” can see totally different things, even things that are not actually there!) arising from this research about the cognitive differences between experts and novices stimulated others to conduct such studies (Charness, 1976, 1979, 1981; Chi, 1978; Chi, Feltovich, & Glaser, 1981; Elstein, Shulman, & Sprafka, 1978; Larkin, McDermott, Simon, & Simon, 1980), and the rest, as they say, is history. The existence of this Cambridge Handbook is its own best evidence for the subsequent development and tremendous expansion of the field of “Expertise Studies” into its current myriad forms.

It is interesting to think about whether a field of expertise studies could have emerged at all—and if so, what it could possibly have looked like—if alternatives to behaviorism had not emerged. For instance, would we have discovered that experts do not just complete tasks and solve problems faster and better than novices, but often attain their solutions in qualitatively different ways? Would we have discovered that experts frequently spend a greater proportion of their time in initial problem evaluation compared to novices (e.g., Glaser & Chi, 1988, regarding “Experts spend a great deal of time analyzing a problem qualitatively”; Lesgold et al., 1988; see also Kellogg’s Chapter 22 on planning by professional writers and Noice & Noice’s Chapter 28 on the deep encoding by professional actors as they study their lines)?

We will, of course, never know, but there was considerable interest in complex thought processes among some of the behaviorists. For example, John B. Watson (1920) was the first investigator to study problem solving by instructing a participant to think aloud while the participant figured out the function of an object (Ericsson, Chapter 13). Neo-behaviorists, such as Berlyne (1965), proposed stimulus-response accounts for complex goal-directed thought and cognitive development. Today, behavior analysts recommend the collection of think-aloud protocols to better understand complex performance (Austin, 2000). Given the broad divide in the theoretical mechanisms used by cognitive and behavioral researchers, it is interesting that researchers are converging on methods of collecting observable process indicators and have mutual interest in large, reproducible differences in performance.
The last peg in the story of expertise studies that we consider is education and educational psychology. There are at least two dimensions in the evolution of education that are related to expertise studies, and that we have also seen in the other influences we have considered. First, like psychology, educational theory and practice was under the influence of behaviorism in and around the mid century (Skinner, 1960; Watson, 1913). Both learning and teaching centered around establishing appropriate stimulus-response connections. “Programmed learning” and “teaching machines” were in vogue. A representative example is the landmark volume co-edited by Robert Glaser (Lumsdaine & Glaser, 1960), who would go on to play a central role in newer incarnations of educational and psychological theory and practice. Essentially, a teaching machine, in doing programmed learning, would present questions or problems to learners, one by one, and depending on the student’s response either reinforce a correct response or note an incorrect one (and perhaps also provide some remedial guidance). This process was believed to establish stable connections between problematic situations and appropriate situational responses. What would expertise look like under such a worldview? It is interesting in this regard to examine a statement about this made by one of Behaviorism’s founders:

Mathematical behavior is usually regarded not as a repertoire of responses involving numbers and numerical operations, but as evidence of mathematical ability or the exercise of the power of reason. It is true that the techniques which are emerging from the experimental study of learning are not designed to “develop the mind” or to further some vague “understanding” of mathematical relationships. They are designed, on the contrary, to establish the very behaviors which are taken to be evidences of such mental states or processes. (Skinner, 1960, pp. 111)

In this view, it seems expertise would be a matter of responding well in challenging situations. Although modern views of expertise retain this criterion of superior performance, there is also considerable interest and theorizing about mediating processes and structures that support, and can be developed to produce, these superior performances (see later sections in this chapter on “Expertise Involves Larger and More Integrated Cognitive Units”; and “Expertise Involves Functional, Abstracted Representations of Presented Information”). Interestingly, however, current theorizing about the critical role of deliberate practice in the development of expertise emphasizes mechanisms not incompatible with these earlier theories, in particular the need for clear goals, repeated practice experiences, and the vital role of feedback about the quality of attempts (Ericsson, Krampe, & Tesch-Römer, 1993). In addition, it is possible that discoveries from behaviorist research about different “schedules of reinforcement” (e.g., Ferster & Skinner, 1957), and their relation to sustaining motivation and effort over long periods of time, might contribute to our understanding of how some people manage to persevere through the very long periods of practice and experience, involving both successes and inevitably many failures, that we now know are so essential to the development of expert levels of skill. How to scaffold sustained, consistent, purposeful effort, over very long periods of time and despite inevitable setbacks, appears at this time to be one of the great puzzles to be solved in developing a science of human excellence (see Hunt, Chapter 3, for a discussion).

With the emergence of the cognitive turn in psychology and educational psychology, a new role for expertise studies also emerged. Expert cognition was conceived as the “goal state” for education, the criterion for what the successful educational process should produce, as well as a measure by which to assess its progress. In this regard, advanced methods have now been developed for eliciting and representing the knowledge of experts (see Hoffman & Lintern, Chapter 12) and for observing and describing experts’ work practices in natural settings (see Clancey, Chapter 8). Novice cognition (as well as that of various levels of intermediates) could serve as “initial states,” as models of the starting place for the educational process. In a sort of means-ends
analysis, the job of education was to determine the kinds of operations that could transform the initial conditions into the desired more expertlike ones (Glaser, 1976). Although it is tempting to believe that upon knowing how the expert does something, one might be able to “teach” this to novices directly, this has not been the case (e.g., Klein & Hoffman, 1993). Expertise is a long-term developmental process, resulting from rich instrumental experiences in the world and extensive practice. These cannot simply be handed to someone (see the later section in this chapter on “Simple Experience Is Not Sufficient for the Development of Expertise”).

One venue in which expertise as “goal state” has gained considerable use is intelligent computer-based education, for example, “intelligent tutoring systems.” (e.g., Clancey & Letsinger, 1984; Forbus & Feltovich, 2001; Sleeman & Brown, 1982). Such systems often utilize an “expert model,” a representation of expert competence in a task, and a “student model,” a representation of the learner’s pertinent current understanding. Discrepancy between the two often drives what instructional intervention is engaged next. Another educational approach is to build tools for enhancing and accelerating experience (e.g., Klein & Hoffman, 1993; Spiro, Collins, Thota, & Feltovich, 2003), and this is closely related to methods for analyzing the representative tasks to be mastered (see Schraagen, Chapter 11).

Some early research on the difference between experts and novices led directly to the creation of new methods of instruction. This is particularly true in medical education, where early expert-novice studies (Barrows, Feightner, Neufeld, & Norman, 1978; Elstein, Shulman, & Sprafka, 1978) led to the creation of “problem-based learning” (Barrows & Tamblyn, 1980). Over a long period of time, PBL (and variants) has come to pervade medical education, as well as making significant inroads into all types of education, including K-12, university, and every sort of professional education (see Ward, Williams, & Hancock, Chapter 14, for a review of the use of simulation in training).

A second theme related to expertise studies that also appears in education, as well as in the other contributors we have discussed, is related to weak and strong methods (Amirault & Branson, Chapter 5). As long as there has been education, there has been controversy about what constitutes an educated person, what such a person should know and be able to do, and how to bring such a person about. Examination of the history of education as it relates to expertise (Amirault & Branson, Chapter 5) reveals the ebb and flow between understanding the object of education (expertise) to be the generalist (sound reasoning, broad knowledge, critical thinking) or the specialist (one who has undergone a great amount of training and experience in a limited domain of activity and has acquired a vast knowledge base specifically tailored for that activity). As with the development of artificial intelligence, our modern educational and psychological conception of expertise seems to favor the specialist and specialized skills, honed over many years of extensive training and deliberate practice (Ericsson, Chapter, 38). The notion of an “expert generalist” is difficult to capture within the current explanatory systems in expertise studies (e.g., Feltovich, Spiro, & Coulson, 1997; see also the discussion of the preparation for creative contributions by Weisberg, Chapter 42).

**Toward Generalizable Characteristics of Expertise and Their Theoretical Mechanisms**

From the kinds of beginnings just discussed, expertise studies have become a large and active field. Fortunately, periodic volumes have served to capture its state of development over time (Anderson, 1981; Bloom, 1985; Chase, 1973; Chi, Glaser, & Farr, 1988; Clancey & Shortliffe, 1984; Ericsson, 1996a; Ericsson & Smith, 1991a; Feltovich, Ford, & Hoffman, 1997; Hoffman, 1992; Starkes & Allard, 1995; Starkes & Ericsson, 2003).

The remainder of the current chapter attempts to crystallize the classic and
enduring findings from the study of expertise. It will draw on generalizable characteristics of expertise identified in earlier reviews (Glaser & Chi, 1988; Chi, Chapter 2) and discuss them and other aspects in the light of the pioneering research that uncovered them. We will also discuss the original theoretical accounts for these findings. However, where pertinent, we will also present more recent challenges and extensions to these classic accounts, including pertinent findings and theoretical treatments reviewed in the chapters of this handbook.

Expertise Is Limited in Its Scope and Elite Performance Does Not Transfer

There is a general belief that talented people display superior performance in a wide range of activities, such as having superior athletic ability and superior mental abilities. However, if we restrict the claims to individuals who can perform at very high levels in a domain, then it is clear that people hardly ever reach an elite level in more than a single domain of activity (Ericsson & Lehmann, 1996). This has proven to be one of the most enduring findings in the study of expertise (see Glaser & Chi, 1988, Characteristic 1). There is little transfer from high-level proficiency in one domain to proficiency in other domains—even when the domains seem, intuitively, very similar.

For example, in tasks similar to those used in the Simon and Chase chessboard studies, Eisenstadt and Kareev (1979) studied the memory for brief displays for expert GO and Gomoko players. Even though these two games are played on the same board and use the same pieces, GO players showed quite poor performance on Gomoko displays, and vice versa. In tasks involving political science, for example, devising plans for increasing crop production in the Soviet Union, Voss and colleagues (Voss, Greene, Post, & Penner, 1983; Voss, Tyler, & Yengo, 1983) found that experts in chemistry (chemistry professors) performed very much like novices in political science, in comparison to political science experts (see Voss & Wiley, Chapter 33, and Endsley, Chapter 36, for more recent examples). Task specificity is also characteristic of expertise involving perceptual-motor skills (e.g., Fitts & Posner, 1967; Rosenbaum, Augustyn, Cohen, & Jax, Chapter 29), as exemplified in many chapters in this handbook, but in particular in perceptual diagnosis and surgery (Norman, Eva, Brooks, & Hamstra, Chapter 19), sports (Hodges, Starkes, & MacMahon, Chapter 27), and music (Lehmann & Gruber, Chapter 26).

Some of the most solid early evidence for specificity in expertise came from expert-novice difference studies in medicine, investigating the clinical reasoning of practitioners (Barrows et al., 1978; Elstein et al., 1978). These studies showed that the same physician can demonstrate widely different profiles of competence, depending on his or her particular experiential history with different types of cases. Indeed, in modern medical education, where assessment of clinical skill is often evaluated by performance on real or simulated cases, it has been found that because of the case-specificity of clinical skill, a large number of cases (on the order of fourteen to eighteen) are needed to achieve an acceptably reliable assessment of skill (Petrusa, 2002; Norman et al., Chapter 19).

Knowledge and Content Matter Are Important to Expertise

In and around the late 1960s and the 1970s, maintaining a traditional distinction between domain-specific skills and general cognitive abilities was becoming less tenable. In research studies, knowledge was no longer seen as a "nuisance variable" but as a dominant source of variance in many human tasks. In particular, Newell and Simon (1972) found that problem solving and skilled performance in a given domain were primarily influenced by domain-specific acquired patterns and associated actions. Domain-specific skills and knowledge were also found to influence even basic cognitive abilities. For example, Glaser and others (Pellegrino & Glaser, 1982a, 1982b) investigated basic foundations of intelligence, including induction, and found evidence that even these were strongly influenced by a person's
knowledge in the operative domain (for example, a person’s conceptual knowledge about numbers in number analogy and number-series tasks).

Acquired knowledge in a domain was found to be associated with changes in fundamental types of cognitive processing. For example, drawing on the expert-novice paradigm, Chi (1978) compared experienced chess-playing children with other children in their performance on memory and learning tasks related to chess. The differences in experience, knowledge, and skill in chess produced differences, in favor of the chess players, in such basic learning processes as the spontaneous use of memory strategies (like grouping and rehearsal), the ability to use such strategies even under experimental prompting, and the amount of information that could be held in short-term memory (Chi, 1978).

Voss and colleagues (Chiesi, Spilich, & Voss, 1979; Spilich, Vesonder, Chiesi, & Voss, 1979) extended this kind of research into other forms of learning. Studying high- and low-knowledge individuals with regard to the game of baseball, they found that, compared to the low-knowledge individuals, high-knowledge ones exhibited superior learning for materials from that and only that particular domain. In particular, high-knowledge individuals had greater recognition and recall memory for new material, could make useful inferences from smaller amounts of partial information, and were better able to integrate new material within a coherent and interconnected framework (organized, for instance, under a common goal structure).

Some studies showed reasoning itself to be dependent on knowledge. Wason and Johnson-Laird (1972) presented evidence that individuals perform poorly in testing the implications of logical inference rules (e.g., if p then q) when the rules are stated abstractly. Performance greatly improves for concrete instances of the same rules (e.g., “every time I go to Manchester, I go by train”). Rumelhart (1979), in an extension of this work, found that nearly five times as many participants were able to test correctly the implications of a simple, single-conditional logical expression when it was stated in terms of a realistic setting (e.g., a work setting: “every purchase over thirty dollars must be approved by the regional manager”) versus when the expression was stated in an understandable but less meaningful form (e.g., “every card with a vowel on the front must have an integer on the back”).

These kinds of studies in the psychology of learning and reasoning were mirrored by developments within artificial intelligence. There was an evolution from systems in which knowledge (declarative) and reasoning (procedural) were clearly separated, to systems in which these components were indistinct or at least strongly interacted. For example, early computer systems, such as Green’s QA3 (Green, 1969) and Quillian’s TLC (Quillian, 1969), utilized databases of declarative knowledge and a few general-purpose reasoning algorithms for operating on those knowledge bases. Such systems were progressively supplanted by ones in which the separation between knowledge and reasoning was not nearly as distinct, and in which general reasoning algorithms gave way to more narrowly applicable reasoning strategies, embedded in procedures for operating within specific domains of knowledge (e.g., Norman, Rumelhart, & LNR, 1979; Sacerdoti, 1977; VanLehn & Seely-Brown, 1979; Winograd, 1975).

It was within this kind of context that studies of expertise and expert-novice differences, along with the growth of knowledge-intensive “expert systems” in artificial intelligence (e.g., Shortliffe, 1976; Buchanan, Davis, & Feigenbaum, Chapter 6), began also to emphasize the criticality of knowledge. This was evident in the progression in AI from weak to strong methods, and within psychology in the growing recognition of the role in expertise of such knowledge-based features as perceptual chunking, knowledge organization, knowledge differentiation, and effective perceptual-knowledge coupling.

This research clearly rejects the classical views on human cognition, in which general abilities such as learning, reasoning,
problem solving, and concept formation correspond to capacities and abilities that can be studied independently of the content domains. In fact, inspired by the pioneering work by Ebbinghaus (1885/1964) on memory for nonsense syllables, most laboratory research utilized stimulus materials for which the prior experience of participants was minimized, in order to allow investigators to study the cognitive processes of learning, reasoning, and problem solving in their “purest” forms. This kind of research, some examples of which were discussed earlier in this section, showed that participants, when confronted with unfamiliar materials in laboratory tasks, demonstrated surprisingly poor performance. In contrast, when tested with materials and tasks from familiar domains of everyday activity, people exhibited effective reasoning, learning, and problem solving. Similarly, the performance of experts is superior to novices and less-skilled individuals primarily for tasks that are representative of their typical activities in their domain of expertise – the domain specificity of expertise (see the earlier section “Expertise Is Limited”).

In the expert-performance approach to expertise, researchers attempt to identify those tasks that best capture the essence of expert performance in the corresponding domain, and then standardize representative tasks that can be presented to experts and novices. By having experts repeatedly perform these types of tasks, experimenters can identify, with experimental and process-tracing techniques, those complex mechanisms that mediate their superior performance (Ericsson, Chapter 13 and Chapter 38). The experts’ superior performance on tasks related to their domain of expertise can be described by psychometric factors (expert reasoning and expert working memory) that differ from those general ability factors used to describe the performance of novices (Horn & Masunaga, Chapter 34, and see Ackerman & Beier, Chapter 9, for a review of individual differences as a function of level of expertise). In short, knowledge matters (Steier & Mitchell, 1996).

**Expertise Involves Larger and More Integrated Cognitive Units**

With increased experience and practice, most people cognitively organize the perceptually available information in their working environment into larger units. This is a classic and one of the best-established phenomenon in expertise (Glaser & Chi, 1988, Characteristic 2). It is supported by a long line of research, but was first discovered in the game of chess (see also Gobet & Charness, Chapter 30).

In the 1960s and early 1970s, de Groot (1965) and Chase and Simon (1973a, 1973b) studied master-level and less-accomplished chess players. In the basic experimental task, participants were shown a chess board with pieces representing game positions from real games. Participants were shown the positions for only five seconds, and they were then asked to reproduce the positions they had seen.

After this brief glance, an expert was able to reproduce much more of the configuration than a novice. In the studies by Chase and Simon (1973a, 1973b) noted earlier, the expert recalled four to five times the number of pieces recalled by the novice. In the similar studies by de Groot, the recall performance by world-class players was nearly perfect (for 25-piece boards). In contrast, novices were able to reproduce about five pieces, or about the number of items that can be maintained in short-term memory exclusively by rehearsal.

The original, classical explanation by Chase and Simon (Simon & Chase, 1973; Chase & Simon, 1973a, 1973b) for expert superiority involved “chunking” in perception and memory. With experience, experts acquire a large “vocabulary,” or memory store, of *board patterns involving groups of pieces*, or what were called chunks. A chunk is a perceptual or memory structure that bonds a number of more elementary units into a larger organization (e.g., the individual letters “c”, “a,” and “r” into the word “car”). When experts see a chess position from a real game, they are able to rapidly recognize such familiar patterns. They can then associate
these patterns with moves stored in memory that have proven to be good moves in the past. Novices do not have enough exposure to game configurations to have developed many of these kinds of patterns. Hence they deal with the board in a piece-by-piece manner. Similarly, when experts are presented with chess boards composed of randomly placed pieces that do not enable the experts to take advantage of established patterns, their advantage over novices for random configurations amounts to only a few additional pieces.

These basic phenomena attributed to chunking were replicated many times, in chess but also in other fields (e.g., the games of bridge, Engle & Bukstel, 1978; GO, Reitman, 1976; and electronics, Egan & Schwartz, 1979). In many such studies, it is the chunk size that is larger for experts. Both the novice and the expert are constrained by the same limitations of short-term (or working) memory (Cowan, Chen, & Rouder, 2004; Miller 1956). However, expert chunks are larger. A chess novice sees a number of independent chess pieces; the expert recognizes about the same number of larger units. For example, one chunk of chess pieces for an expert might be a “king defense configuration,” composed of a number of individual chess pieces.

As we have just discussed, it was originally believed that experts develop larger chunks and that these enable the expert to functionally expand the size of short-term or working memory. However, in the mid-1970s, Charness (1976) showed that expert chess players do not rely on a transient short-term memory for storage of briefly presented chess positions. In fact, they are able to recall positions, even after the contents of their short-term memory have been completely disrupted by an interfering activity. Subsequent research has shown that chess experts have acquired memory skills that enable them to encode chess positions in long-term working memory (LTWM, Ericsson & Kintsch, 1995). The encoding and storage of the chess positions in LTWM allow experts to recall presented chess positions after disruptions of short-term memory, as well as being able to recall multiple chess boards presented in rapid succession (see Ericsson, Chapter 13, and Gobet & Charness, Chapter 30, for an extended discussion of new theoretical mechanisms accounting for the experts’ expanded working memory). The experts’ superior ability to encode representative information from their domain of expertise and store it in long-term memory, such that they can efficiently retrieve meaningful relations, provides an alternative to the original account of superior memory in terms of larger chunks stored in STM. There is another, similar characteristic of expertise. It has to do with the nature and organization of the perceptual encoding and memory structures experts develop and use. This is discussed next.

**Expertise Involves Functional, Abstracted Representations of Presented Information**

Some studies, utilizing methods similar to the Simon and Chase chessboard paradigm, examined the nature of expert and novice cognitive units, such as chunks or other knowledge structures. Chase and Simon (1973a, 1973b) themselves analyzed the characteristics of the chess pieces their experts grouped together as they reproduced a chess position after a brief presentation. Expert configurations of chess pieces were based largely on strategic aspects of the game, for example, configurations representing elements of threat or opportunity. It was not clear how novice units were organized. Glaser and Chi (1988) identified a related general characteristic, namely, that “Experts see and represent a problem in their domain at a deeper (more principled) level than novices; novices tend to represent a problem at a superficial level” (p. xviii). Our characterization for expert representations, “functional and abstracted” as elaborated next, simply seeks to provide a bit more insight into the nature of “deep” (see Chi, Chapter 10, for a review of research on assessments of experts’ cognitive representations).

Early studies involving bridge (Charness, 1979, Engle & Bukstel, 1978) and electronics
studies of expertise from psychological perspectives

(Egan & Schwartz, 1979), patterned after the Chase and Simon procedure, showed similar results. In the bridge studies, experts and novices were briefly presented depictions of four-handed bridge deals, and they were required to reproduce these deals. Experts reproduced the cards by suit, across hands. They remembered cards of the same suit from three hands and inferred the fourth; this is an organization useful in playing the game of bridge. Novices recalled the cards by order of card rank within hands, an organization not useful to supporting strategic aspects of the game. In electronics, subjects were shown an electronic circuit diagram, which they were then to reproduce. Experts grouped individual diagram components into major electronic components (e.g., amplifiers, filters, rectifiers). Novice organization was based largely on the spatial proximity of symbols appearing in the diagram.

Similar results have been shown from yet other fields, using somewhat different methodologies that compared the performance of groups of adults who differ in their knowledge about a given domain. For example, Voss and colleagues (Spilich et al., 1979) studied ardent baseball fans and more casual baseball observers. Participants were presented a colorful description of a half-inning of baseball and were then asked to recall the half-inning. Expert recall was structured by major goal-related sequences of the game, such as advancing runners, scoring runs, and preventing scoring. Novices’ recall contained less integral components, for example, observations about the weather and the crowd mood. Novice recall did not capture basic game-advancing, sequential activity nearly as well. More recent research on fans that differ in their knowledge about soccer and baseball has found that comprehension and memory for texts describing games from these sports is more influenced by relevant knowledge than by verbal IQ scores (see Hambrick & Engle, 2002, for a recent study and a review of earlier work).

Two early studies of computer programming produced similar results. McKeithen, Reitman, Reuter, and Hirtle (1981) presented a list of 21 commands in the ALGOL language to ALGOL experts, students after one ALGOL course, and students at the beginning of an ALGOL course. Participants were given 25 recall trials after they initially learned the list. The organization of the recalled items by pre-ALGOL students was by surface features of commands (e.g., commands with the same beginning letter or same length of command name) and groups of commands forming natural language segments (e.g., “STRING IS NULL BITS”) that have no conceptual meaning within the language. Experts, in contrast, grouped commands that formed mini ALGOL algorithms (e.g., formation of loops) or constituted types of ALGOL data structures. Students, after an ALGOL course, produced groupings that were a mixture of surface-related and meaningful ALGOL organizations.

In a similar study, Adelson (1981) presented a list of programming commands, constituting three intact computer programs, scrambled together and out of order, to expert and novice programmers. Participants were required to recall the list. Over recall trials, experts reconstructed the original three algorithms. The organization of novice recall was by syntactic similarities in individual command statements, regardless of the embedded source algorithms. Sonnenstag, Niessen, and Volmer (Chapter 21) provide a review of the more recent research on knowledge representations and superior performance of software experts.

Other pertinent findings came from early work in physics (Chi et al., 1981) and medicine (Feltovich, Johnson, Moller, & Swanson, 1984; Johnson et al., 1981). In the basic task from the physics study, problems from chapters in an introductory physics text were placed on individual cards. Expert (professors and advanced graduate students) and novice (college students after their first mechanics course) physics problem solvers sorted the cards into groups of problems they would “solve in a similar manner.” The finding was that experts created groups based on the major physics principles (e.g., conservation and force laws) applicable in
the problems’ solutions. Novice groupings were organized by salient objects (e.g., springs, inclined planes) and features contained in the problem statement itself. Similarly, in studies of expert and novice diagnoses within a subspecialty of medicine, expert diagnosticians organized diagnostic hypotheses according to the major pathophysiological issue relevant in a case (i.e., constituting the ‘Logical Competitor Set’ of reasonable alternatives for the case, e.g., lesions involving right-sided heart volume overload), whereas novice hypotheses were more isolated and more dependent on particular patient cues.

Zeitz (1997) has reviewed these and more recent studies of this type, investigating what she calls experts’ use of “Moderately Abstracted Conceptual Representations” (MACRs), which are representational abstractions of the type just discussed. She proposes numerous ways in which such abstraction aids the efficient utilization of knowledge and reasoning by experts. These include: (a) the role of abstracted representations in retrieving appropriate material from memory (e.g., Chi et al., 1981); (b) the schematic nature of MACRs in integrating information and revealing what information is important; (c) providing guidance for a line of action and supporting justification for such a line of approach (e.g., Phelps & Shanteau, 1978; Schmidt et al., 1989; Voss et al., 1983); (d) aiding productive analogical reasoning (e.g., Gentner, 1988); and (e) providing abstract representations that support experts’ reasoning and evaluation of diagnostic alternatives (e.g., Patel, Arocha, & Kaufman, 1994).

The functional nature of experts’ representations extends to entire activities or events. Ericsson and Kintsch (1995) proposed that experts acquire skills for encoding new relevant information in LTWM to allow direct access when it is relevant and to support the continual updating of a mental model of the current situation – akin to the situational models created by readers when they read books (see Endsley, Chapter 36, on the expert’s superior ability to monitor the current situation – “situational awareness”). This general theoretical framework can account for the slow acquisition of abstract representations that support planning, reasoning, monitoring, and evaluation (Ericsson, Patel, & Kintsch, 2000). For example, studies of expert fire fighters have shown that experts interpret any scene of a fire dynamically, in terms of what likely preceded it and how it will likely evolve. This kind of understanding supports efforts to intervene in the fire. Novices interpret these scenes in terms of perceptually salient characteristics, for example, color and intensity (Klein, 1998, and see Ross, Shafer, & Klein, Chapter 23). Studies of expert surgeons have shown that some actions within a surgery have no real value for immediate purposes, but are made in order to make some later move more efficient or effective (Koschmann, LeBaron, Goodwin, & Feltovich, 2001). The research on expert chess players shows consistent evidence for extensive planning and evaluation of consequences of alternative move sequences (See Ericsson, Chapter 13, and Gobet & Charness, Chapter 30). Furthermore, there is considerable evidence pertaining to experts’ elaborated encoding of the current situation, such as in situational awareness (Endsely, Chapter 36), mental models (Durso & Datel, Chapter 20), and LTWM (Noice & Noice, Chapter, 28).

In summary, research conducted in the last thirty or so years indicates that expert performers acquire skills to develop complex representations that allow them immediate and integrated access to information and knowledge relevant to the demands of action in current situations and tasks. These acquired skills can account for their superior memory performance when they are given a task, such as recalling a briefly presented chess position, as in the studies by Chase and Simon (1973a, 1973b). Novices, on the other hand, lack such knowledge and associated representations and skills, and thus perform these tasks with the only knowledge and skills they have available. They try to impose organization and meaningful relations, but their attempts are piecemeal and less relevant to effectively
functioning in the task domain, organized, for example, by items named in a situation, current salient features, proximity of entities to others, or superficial analogies.

**Expertise Involves Automated Basic Strokes**

Most people considered to be experts are individuals with extreme amounts of practice on a circumscribed set of tasks in their work environment. For example, some expert radiologists estimated they had analyzed more than half a million radiographs (X-rays) in their careers (Lesgold et al., 1988). Such experience, appropriately conducted, can yield effective, major behavioral and brain changes (Hill & Schneider, Chapter 37).

Research on the effects of practice has found that the character of cognitive operations changes after even a couple of hours of practice on a typical laboratory task. Operations that are initially slow, serial, and demand conscious attention become fast, less deliberate, and can run in parallel with other processes (Schneider & Shiffrin, 1977). With enough practice, one can learn how do several tasks at the same time. Behavioral studies of skill acquisition have demonstrated that automaticity is central to the development of expertise, and practice is the means to automaticity (Posner & Snyder, 1975, see also Proctor & Vu, Chapter 15). Through the act of practice (with appropriate feedback, monitoring, etc.), the character of cognitive operations changes in a manner that (a) improves the speed of the operations, (b) improves the smoothness of the operations, and (c) reduces the cognitive demands of the operations, thus releasing cognitive (e.g., attentional) resources for other (often higher) functions (e.g., planning, self-monitoring; see also Endsley Chapter 36). Automatic processes seem resistant to disruption by reduced cognitive capacity and, to a limited degree, are largely resource insensitive (Schneider & Fisk, 1982). Interestingly, fMRI studies have demonstrated that shifts to automaticity reveal a shift (decrease) in activity in a certain part of the brain, but not a shift in anatomical loci (Jansma, Ramsey, Slagter, & Kahn, 2001; Hill & Schneider, Chapter 37).

There are many examples in the early expertise-related literature of the effects of practice on dual-task performance of experts. For example, expert typists can type and recite nursery rhymes at the same time (Shaffer, 1975). Skilled abacus operators can answer routine questions (“What is your favorite food?”) without loss of accuracy or speed in working with the abacus (Hatano, Miyake, & Binks, 1977). After six weeks of practice (one hour per day), college students could read unfamiliar text while simultaneously copying words read by an experimenter, without decrement in reading speed or comprehension (Spelke, Hirst, & Neisser, 1976).

Automaticity is important to expertise. It appears it has at least two main functions. The first has to do with the relationship between fundamental and higher-order cognitive skills, and the second has to do with the interaction between automaticity of processes and usability of available knowledge.

With regard to the first, in complex skills with many different cognitive components, it appears that some of the more basic ones (e.g., fundamental decoding, encoding of input) must be automated if higher-level skills such as reasoning, comprehension, inference, monitoring, and integration are ever to be proficient (e.g., Logan, 1985; Endsley, Chapter 36). For example, in a longitudinal study, Lesgold and Resnick (1982) followed the same group of children from their initial exposure to reading in kindergarten through third and fourth grade. They found, for example, that if basic reading skills do not become automated, such as the decoding and encoding of letters and words, comprehension skills will not substantially develop. Furthermore, the relationship seems to be causal; that is, speed increases in word skills predict comprehension increases later on, whereas increases in comprehension do not predict increases in word facility. However, subsequent pertinent research has accentuated the complex-nature of the relationship...
between automated basic processes and
higher-order deliberate ones and point to
the need for continued research (Hill &
Schneider, Chapter 37).

There is also a possible interaction
between automaticity of processes and the
usability of available knowledge. Investiga-
tors (e.g., Feltovich et al., 1984; Jeffries
et al., 1981) have suggested that a major
limitation of novices is their inability to
access knowledge in relevant situations, even
when they can retrieve the same knowledge
when explicitly cued by the experimenter.
Problems in knowledge usability may be
associated with overload or inefficiency in
using working (or short-term) memory. The
usable knowledge of experts may, in turn,
result from the subordination of many task
components to automatic processing, which
increases capability for controlled manage-
ment of memory and knowledge application

An alternative proposal about usability
of knowledge has subsequently been made
by Ericsson and Kintsch (1995), in which
experts acquire skills that are designed to
encode relevant information in long-term
memory (LTM) in a manner that allows
automatic retrieval from LTM when later
needed, as indicated by subsequent activa-
tion of certain combinations of cues in
attention. They argued that experts acquire
LTWM skills that enable them, when they
encounter new information (such as a
new symptom during an interview with a
patient), to encode the relevant associations
such that when yet other related information
is encountered (such as subsequent informa-
tion reported by the patient), the expert
will automatically access relevant aspects
of the earlier information to guide encod-
ing and reasoning. The key constraint for
skilled encoding in LTM is that the expert
be able to anticipate potential future con-
texts where the encountered information
might become relevant. Only then will the
expert be able to encode encountered informa-
tion in LTWM in such a way that its
future relevance is anticipated and the re-
levant pieces of information can be automat-
ically activated when the subsequent rele-
vant contexts are encountered. In this model
of the experts’ working memory storage in
LTM, the large capacity of LTM allows the
expert to preserve access to a large body of
relevant information without any need to
actively maintain the information in a lim-
ited general capacity STM (Ericsson, Chap-
ter 13; Gobet & Charness, Chapter 30; Noice
& Noice, Chapter 28; Wilding & Valentine,
Chapter 31).

**Expertise Involves Selective Access
of Relevant Information**

Within the classical expertise framework
based on chunking, questions about access
to task-relevant information are important
issues, and a critical aspect of intelligence
(Sternberg, 1984). Given the functional
nature of expert representations, how are
they properly engaged in the context of solv-
ing a problem? To what kind of problem fea-
tures do experts attend? How are these fea-
tures “linked up” to the significant concepts
in memory? In a sense, having a trace laid
down in memory is not a sufficient condi-
tion for use. Extant traces must be accessed
and important non-extant traces must be
inferred or otherwise computed.

This characteristic of expertise addresses
the critical problem of accessing knowledge
structures. This development overcomes (at
least) two difficulties for expertise as a “big
switch” (Newell, 1973) between the recog-
nition of familiar events and application
of experience associated with those events
(see also Ross, Shafer, & Klein, Chapter 23,
“recognition-primed decision making). The
first of these is related to the variability
in events; one cannot “step into the same
river twice.” The useful utilization of events
as familiar requires a degree of appropriate
abstraction, both in the event features
utilized and in the memory organization
imposed on the memory models them-
seves. The former adaptation is reflected in
expert utilization of abstracted features for
problem classification, features whose loci
in a problem statement are not apparent
(Chi et al., 1981). The latter adaptation is
reflected in the development of hierarchical
organizations, which characterize expert or experienced memory (e.g., Feltovich et al., 1984; Patil, Szolovits, & Schwartz, 1981).

Critical to this characteristic is selectivity. Selectivity is based on the attribution of differential importance or, broadly conceived, a separation of signal from noise either in the features extracted from events or on internal cognitive processes themselves (see also Hill & Schneider, Chapter 37). Selectivity, as a means of task adaptation, is assumed to be forced on the human based on their limited cognitive capacity. With regard to events, selectivity involves the abstraction of invariances of the discriminating cues that define types of situations or are otherwise integral to a task. Expertise, then, involves learning which information is most useful and which is tangential or superfluous (e.g., Chi et al., 1981; Hinsley et al., 1978; Patel & Groen, 1991; Spilich et al., 1979). In certain types of "stable" environments, the important invariance is well defined and the task is sufficiently constrained so that the mechanisms linking selectivity and performance can be explicated. For example, as consistent with the LTWM hypothesis, skilled typists appear to achieve subordination, usability, and access by developing integrated representations of letters and key presses that facilitate translation between perception and response (Rieger, 2004).

This theme of expertise also reflects the general problem of knowledge inversion; that is, the notion of moving from a concept-centered mode of reasoning to a mode that must somehow scan the problem features for regularities, incorporate abstraction, integrate multiple cues, and accept natural variation in patterns to invoke aspects of the relevant concept. We find this in many fields. For example, medical students acquire much specific "disease-centered" knowledge - given disease X, this is the underlying pathophysiology, these are the variations, and these are the classic manifestations. When faced with a patient, however, they are presented with just the opposite situation: Given a patient, what is the disease? Recent developments in medical education focus on case-oriented learning in which medical students are given early exposure to representative clinical situations. This type of training forces learners to develop mental representations and an LTWM that support medical reasoning under real-time, representative constraints (Norman, Eva, Brooks, & Hamstra, Chapter 19; Ericsson, Chapter 13; Endsley, Chapter 36).

**Expertise Involves Reflection**

Another challenge to the traditional information processing view, with its severe constraints on cognitive capacity, concerns the experts' ability not just to perform effectively but also to be able to reflect on their thought processes and methods (Glaser & Chi, 1988, Characteristic 7 (see also Zimmerman, Chapter 39)). Metacognition is knowledge about one's own knowledge and knowledge about one's own performance (Flavell, 1979). It is what an individual knows about his or her own cognitive processes. Its relevance to expertise is derived, in part, from the observation that experts are graceful in their reasoning process. As Bartlett (1958) notes, "Experts have all the time in the world." There is an element of unencumbered elegance in expert performance, the underpinnings of which are based on the efficient management and control of the adaptive processes. A source for this might be in abstracted layers of control and planning.

The traditional (classical) account of metacognition within the information-processing model is that abstract descriptions of plans and procedures enable an individual to operate on or manipulate problem-solving operations, for example, to modify and adjust them to context. They also provide a general organizational structure that guides and organizes the details of application, so that a general line of reasoning can be maintained despite low-level (detailed) fluctuations and variations. Novice physics problem solvers, in contrast to experts, have no abstract or meta-level descriptions for their basic problem-solving operators, which for them are physics equations (Chi et al., 1981). Rather,
operators are tied directly to problem details, show little modifiability, and can only organize problem-solving activity locally (i.e., at the level of isolated problem components present).

In addition to abstraction in control and planning, there must also be mechanisms for maintaining information to allow efficient back-tracking or starting over when lines of reasoning need to be modified or abandoned. Largely, the traditional view proposes that experts deal with the severe working-memory demands required by backtracking by minimizing the need for it. For example, experts can attempt to withhold decisions until they are sufficiently constrained to restrict the options. In other cases when decisions are under-constrained, experts can rely on abstract solution descriptions and conditions for solution (constraints) that both guide the search for solutions and help eliminate alternatives.

The traditional information processing view has difficulties in accounting for the possibility that experts might be disrupted or otherwise forced to restart their planning. More recent research has shown that experts are far more able to maintain large amounts of information in working memory. For example, chess masters are able to play chess games with a quality that approaches that of normal chess-playing under blindfolded conditions in which perceptual access to chess positions is withheld (for a review see Ericsson et al., 2000; Ericsson & Kintsch, 2000). Chess masters are able to follow multiple games when they are presented move by move and can recall the locations of all pieces with high levels of accuracy. Chess masters are also able to recall a series of different chess positions when they are briefly presented (5 seconds per position). In studies of expert physicians (e.g., Feltovich, Spiro, & Coulson, 1997), it was found that when experts do not know the correct diagnosis for a patient, they often can give a plausible description of the underlying pathophysiology of a disease; that is, they are able to reason at levels that are more fundamental and defensible in terms of the symptoms presented. When novices fail to reach a diagnosis for a patient, their rationale for possible alternatives is generally incompatible with the symptoms presented. Experts fail gracefully; novices crash. Vimla Patel and her colleagues (Groen & Patel, 1988; Patel & Groen, 1991) have found that medical experts are able to explain their diagnoses by showing how the presented symptoms are all explained by the proposed integrated disease state, whereas less advanced medical students have a more piece-meal representation that is less well integrated.

Metacognition, then, is important for people to test their own understanding and partial solutions to a problem. This kind of monitoring prevents blind alleys, errors, and the need for extensive back-up and retraction, thus ensuring overall progress to a goal. In addition, these same kinds of monitoring behaviors are critical throughout the process of acquiring knowledge and skills on which expertise depends. The mental representations developed by aspiring experts have multiple functions. They need to allow efficient and rapid reactions to critical situations, and they need to allow modifiability, mechanisms by which a skilled performer, for instance, adjusts his performance to changed weather conditions, such as a tennis player dealing with rain or wind, or adjusts to unique characteristics of the place of performance, such as musicians adjusting their performance to the acoustics of the music hall. Furthermore, these representations need to be amenable to change so aspiring expert performers can improve aspects and gradually refine their skills and their monitoring representations.

Experts, for the most part, work in the realm of the familiar (familiar for them, not for people in general) and may often be able to generate adequate actions by rapid recognition-based problem solving (Klein, 1998). The same experts are also the individuals called on to address the subtle, complicated, and novel problems of their field. They need to recognize when the task they are facing is not within their normal, routine domain of experience and adjust accordingly (Feltovich et al., 1997); this is just one of many pertinent aspects of
metacognitive activity in the function of expertise.

If the view is maintained that metacognition (in the broadest sense) is enabled by metacognitive knowledge, and metacognitive knowledge is, in fact, “knowledge,” should we not expect it to be subject to the same demands and possess the same properties as “regular” knowledge, albeit in a slightly different context? Evidence exists, for example, that metacognition can be automatic (Reder & Shunn, 1996), thus avoiding Tulving’s (1994) consciousness requirement for metacognitive judgement. There is also indication that metacognitive strategies are explicitly learnable in rather general contexts (Kruger & Dunning, 1999), as well as in special contexts such as reading (Paris & Winograd, 1990) and nursing (Kuiper & Pesut, 2004). Accordingly, metacognitive activities, perhaps in a variety of ways and forms, both explicit and implicit, afford and support the developmental and performance dynamics of expertise.

**Expertise Is an Adaptation**

In this section, we advance an argument that the development of expertise is largely a matter of amassing considerable skills, knowledge, and mechanisms that monitor and control cognitive processes to perform a delimited set of tasks efficiently and effectively. Experts restructure, reorganize, and refine their representation of knowledge and procedures for efficient application to their work-a-day environments (See also Ericsson & Lehmann, 1996). Experts certainly know more, but they also know differently. Expertise is appropriately viewed not as simple (and often short-term) matter of fact or skill acquisition, but rather as a complex construct of adaptations of mind and body, which include substantial self-monitoring and control mechanisms, to task environments in service of representative task goals and activities. As we shall argue, the nature of the adaptations reflects differential demands of the task environment and mediates the performance evidenced by highly skilled individuals. Adaptation matters (Hill and Schneider, Chapter 37).

The classical theory of expertise (Simon & Chase, 1973) focused on the fundamental architectural limits imposed on human information-processing capacities. Early investigators assumed that complex cognition must occur within surprisingly rigidly constrained parameters. Many of these limits are not singular, but are considered collectively as a statement of associated (related) constraints. Furthermore, the architecture underlying these constraints is not specified, other than the fact that it is physical. Thus, the constraints of the architecture could be realized as a symbol system (e.g., Newell & Simon, 1976), perhaps grounded in modalities (Barsalou, 1999; Barsalou et al., 2003), or as a dynamic phase space (e.g., van Gelder & Port, 1995).

In particular, under the traditional theme, three specific reasoning limits are important to explaining performance of typical novices on traditional laboratory tasks (e.g., Prietula & Simon, 1989). First, there is a limit of attention (Shipp, 2004). We can focus on solving only one problem (or making only one decision) at a time when performing an unfamiliar task. However, we sometimes share our attentional resources by shifting rapidly from thinking about a given task to another different task. In addition, perceptual limits on what can be detected with the eye (and the eye-brain) exist, situating the perception in scale bands of size (of objects), time (speed of movement), distance, and spectra. Our perceptual and attention resources have evolved to handle a region of time, a region of space, a region of distance, and a region of spectra. We act in, and react with, a highly constrained perceptual environment, balancing attention and awareness (Lamme, 2003).

Related to this single-mindedness is a limit of working memory. There is a difference between long-term memory, our large, permanent repository for knowledge and working memory, which is much smaller in capacity and restricted to holding information about the particular task at hand, involving multiple components that
mediate between long-term memory and the environment (Baddeley, 2000, 2002). When focusing attention on making a particular decision or solving a particular problem, three types of events occur that are critical for effective reasoning: (1) we seek (and perceive) data from the environment, (2) we bring relevant knowledge to bear from our long-term memory to working memory, and, by reviewing the data in the presence of relevant knowledge retrieved from long-term memory, (3) we draw inferences about what is going on— which may lead to seeking more data and activating more knowledge.

Finally, there is a *limit of long-term memory access*. To what extent we truly forget things is uncertain, so there may not actually be an arbitrary size constraint on this aspect of our long-term memory. That, however, is not the issue. What is certain is that we lose access to (or the power to evoke) the knowledge stored. A typical demonstration of this is the “tip of the tongue” phenomenon—in which you know that you know something, but cannot retrieve it (Brown 1991; Brown & McNeill 1966). Therefore, even though we may scan the right data in an analysis, there is no guarantee that we will be able to trigger the appropriate knowledge in long-term memory to allow us to make correct inferences from those data. In practice, a large part of expert problem solving is being able to access relevant knowledge, at the right time, for use in working memory.

This traditional approach to expertise was founded on the powerful theoretic assumption that experts’ cognitive processes, such as generating, representing, and using knowledge, had to conform to these severe limits. This theory proposed many mechanisms by which experts would be able to functionally adapt to these constraints to produce superior performance. The expert chunking mechanism, for example, permits a vocabulary that is much more robust and complex than the novice can invoke. Although both the expert and the novice have the same working-memory constraints, the expert sees the world in larger and more diverse units. In effect, chunking permits expanding the functional size of working memory and increasing the efficiency of search. This phenomenon has been experimentally demonstrated across a remarkably wide variety of domains.

The role and function of automaticity within expertise is important in this regard also. Automaticity seems to be entwined with functional organization, chunking, and conditions of application. They work in concert to adapt to the demands of the task, under the constraints of both the task and their own capabilities to make appropriate use of our memory. Automaticity, then, is intricately bound with the overall adaptation of the system through knowledge reorganization and refinement.

The general argument is that expert knowledge structures and procedures are reorganized in directions that enable effective application to task demands of a working environment. As we have discussed, most of these changes are adaptations that enable utilization of large amounts of information in the context of limited internal-processing resources (in particular those imposed by the small capacity of short-term or working memory). Grouping or chunking on information structures and procedure components functionally increases the size of working memory and its efficiency. More information can be considered for each “unit” in working memory. Expert selectivity, discrimination, and abstraction (discussed earlier) insure that only the most useful information is thrown into competition for resources. Automaticity is a means of restructuring some procedures so that working memory is largely circumvented, freeing resources for other cognitive chores. It is a tension between high information load and limited internal resources that encourages the development of strategies for the efficient use of knowledge and processing.

This pioneering theory of expertise (Simon & Chase, 1973) has been and remains very influential and has been extended with additional mechanisms to explain experts’ greatly expanded working memory (Gobet & Simon, 1996; Richman, Gobet, Staszewski, & Simon, 1996). At the same
time there have been many arguments raised against the claims that the computational architecture remains fixed and thus presents an invariant constraint on skilled and expert processing.

One of the most general criticisms is that the laboratory – produced empirical evidence for capacity constraints of attention and STM are based on an operational definition of chunks in terms of independent pieces of information, no matter how small or large the individual chunks (see Cowan, 2001, and Ericsson & Kirk, 2001). It is relatively easy to design experimental materials for memory experiments that are made of independent pieces and measure experts' and novices' memory in terms of chunks. However, when one analyzes the information processed by experts when they perform representative tasks in their domain of expertise, then all the heeded and relevant information has relations to the task and other pieces of information. If the encountered information can be encoded and integrated within a model of the current context, then how many independent chunks are stored or maintained in attention and working memory? Similarly, when experts encounter representative tasks situations where beginners perceive several independent tasks, the aspiring experts are able to develop skills and encodings that allow them to integrate the different tasks into a more general task with more diversified demands. More recent research has shown how in laboratory studies, participants performing dual tasks that are believed to contain immutable bottlenecks of processing can, after training, perform them without any observable costs of the dual task (Meyer & Kieras, 1997; Schumacher et al., 2001, but see Proctor & Vu, Chapter 15, for an alternative account). If the definition of chunks and tasks requires independence for imposing limits on information processing, then it seems that the acquisition of expertise entails developing integrated representations of knowledge and coordination of initially separate tasks that make the fundamental information-processing limits inapplicable or substantially attenuated.

The second general criticism of the traditional theory of expertise comes from a rejection of the premise that expertise is an extension of the processes observed in everyday skill acquisition (Fitts & Posner, 1967). According to this model, the acquisition of skill proceeds in stages, and during the first stage people acquire a cognitive representation of the task and how to react in typical situations so they can avoid gross errors. During the subsequent stages, the performance of sequences of actions becomes smoother and more efficient. In the final stage, people are able to perform with a minimal amount of effort, and performance runs essentially automatically without active cognitive control. In an edited book on general theories of expertise (Ericsson & Smith, 1991a), several researchers raised concerns about explaining expertise as an extension of this general model (Ericsson & Smith, 1991b; Holyoak, 1991; and Salthouse, 1991). Ericsson and Smith (1991b) found evidence that experts maintain their ability to control their performance and are able to give detailed accounts of their thought processes that can be validated against other observable performance and process data. Ericsson and Smith reviewed evidence that complex cognitive representations mediate the performance and continued learning by experts, which has been confirmed by subsequent reviews (Ericsson, 1996b, 2003, Chapter 13).

The third and final type of criticism comes from the emerging evidence that extended focused practice has profound effects on, and can influence virtually every aspect of, the human body, such as muscles, nerve systems, heart and circulatory system, and the brain. Several chapters in this handbook review the structural changes resulting from practice, such as Butterworth, Chapter 32, on mathematical calculation; Ericsson, Chapter 38; Lehmann and Gruber, Chapter 26, on music performance; Proctor and Vu, Chapter 15, on adaptations in skill acquisition; and Hill and Schneider, Chapter 37, with an overview of changes in the structure and function of the brain with extended practice and the development of expertise.
Simple Experience Is Not Sufficient for the Development of Expertise

Most everyday skills are relatively easy to acquire, at least to an acceptable level. Adults often learn to drive a car, type, play chess, ski, and play (bad) golf within weeks or months. It is usually possible to explain what an individual needs to know about a given skill, such as rules and procedures, within a few hours (see also Hoffman & Lintern, Chapter 12). Once individuals have learned the underlying structure of the activity and what aspects they must attend to, they often focus on attaining a functional level of performance. This is often attained in less than 50 hours of practice. At this point, an acceptable standard of performance can be generated without much need for more effortful attention and execution of the everyday activity has attained many characteristics of automated performance (Anderson, 1982, 1987; Fitts & Posner, 1967; Shiffrin & Schneider, 1977) and requires only minimal effort.

In their seminal paper, Simon and Chase (1973) pointed to similarities between the decade-long mastery of one’s first language and the need for extended experience to master complex domains of expertise, such as chess and sports. They made a strong argument for a long period of immersion in active participation in activities in the domain, making the claim that even the best chess players needed to spend over ten years studying chess before winning at the international level. The necessity for even the most talented performers to spend ten years working and practicing was later converted into an equivalence, namely, that ten years of experience in a domain made somebody an expert. However, for chess, tennis, and golf, everyone knows examples of excited recreational players who regularly engage in play for years and decades, but who never reach a very skilled level.

Reviews of the relation between the amount of experience and the attained level of performance show consistently that once an acceptable level is attained, there are many domains where performance decreases as a function of the number of years since graduation from the training institution (Ericsson, Chapter 38).

Several research methods have been developed to describe the development paths of expert performers, such as analysis of the historical record of eminent performers (Simonton, Chapter 18), retrospective interviews (see Sosniak, Chapter 16), and diary studies of practice (See Deakin, Côté, & Harvey, Chapter 17). Research with these methods has shown that additional experience appears to make performance less effortful and less demanding, but to improve performance it is necessary to seek out practice activities that allow individuals to work on improving specific aspects, with the help of a teacher and in a protected environment, with opportunities for reflection, exploration of alternatives, and problem solving, as well as repetition with informative feedback.

In this handbook several chapters discuss the effectiveness of this type of deliberate practice in attaining elite and expert levels of performance (Ericsson, Chapter 38; Zimmerman, Chapter 39), in software design (Sonnentag, Niessen, & Volmer, Chapter 21), in training with simulators (Ward, Williams, & Hancock, Chapter 14), in maintaining performance in older experts (Krampe & Charness, Chapter 40), and in creative activities (Weisberg, Chapter 42). Other chapters review evidence on the relationship between deliberate practice and the development of expertise in particular domains, such as professional writing (Kellogg, Chapter 22), music performance (Lehmann & Gruber, Chapter 26), sports (Hodges, Starkesi & MacMohan, Chapter 27), chess (Gobet & Charness, Chapter 30), exceptional memory (Wilding & Valentine, Chapter 31), and mathematical calculation (Butterworth, Chapter 32).

Concluding Remarks

The theoretical interest in expertise and expert performance is based on the
assumption that there are shared psychological constraints on the structure and acquisition of expert performance across different domains. The theory of Simon and Chase (1973) proposed that the invariant limits on information processing and STM severely constrained how expert skill is acquired and proposed a theory based on the accumulation through experience of increasingly complex chunks and pattern-action associations. This theory emphasized the acquired nature of expertise and focused on the long time required to reach elite levels and the learning processes sufficient to gradually accumulate the large body of requisite patterns and knowledge. This view of expertise offered the hope that it would be possible to extract the accumulated knowledge and rules of experts and then use this knowledge to more efficiently train future experts and, thus, reduce the decade or more of experience and training required for elite performance. Efforts were made even to encode the extracted knowledge in computer models and to build expert systems that could duplicate the performance of the experts (Bachanan et al., Chapter 6).

Subsequent research on extended training revealed that it is possible to acquire skills that effectively alter or, at least, circumvent the processing limits of attention and working memory. Studies of expertise focused initially on the expert’s representation and memory for knowledge. As research started to examine and model experts’ superior performance on representative tasks, it became clear that their complex representations and mechanisms that mediate performance could not be acquired by mere experience (Ericsson, Chapter 38). Research on what enabled some individuals to reach expert performance, rather than mediocre achievement, revealed that expert and elite performers seek out teachers and engage in specially designed training activities (deliberate practice). The future expert performers need to acquire representations and mechanisms that allow them to monitor, control, and evaluate their own performance, so they can gradually modify their own mechanisms while engaging in training tasks that provide feedback on performance, as well as opportunities for repetition and gradual refinement.

The discovery of the complex structure of the mechanisms that execute expert performance and mediate its continued improvement has had positive and negative implications. On the negative side, it has pretty much dispelled the hope that expert performance can easily be captured and that the decade-long training to become an expert can be dramatically reduced. All the paths to expert performance appear to require substantial extended effortful practice. Effortless mastery of expertise, magical bullets involving training machines, and dramatic shortcuts, are just myths. They cannot explain the acquisition of the mechanisms and adaptations that mediate skilled and expert performance. Even more important, the insufficiency of the traditional school system is becoming apparent. It is not reasonable to teach students knowledge and rules about a domain, such as programming, medicine, and economics, and then expect them to be able to convert this material into effective professional skills by additional experience in the pertinent domain. Schools need to help students acquire the skills and mechanisms for basic mastery in the domain, and then allow them gradually to take over control of the learning of their professional skills by designing deliberate practice activities that produce continued improvement.

On the positive side, the discovery of effective training methods for acquiring complex cognitive mechanisms has allowed investigators to propose types of training that appear to allow individuals to acquire levels of performance that were previously thought to be unobtainable, except for the elite group of innately talented. The study of the development of expert performers provides observable paths for how they modified or circumvented different types of psychological and physiological constraints. It should be possible for one type of expert in one domain, such as surgery, to learn from how other experts in music or sports, for instance, have designed successful training procedures for mastering various aspects of perceptual-motor procedures, and to learn the amount of practice needed to reach
specified levels of mastery. If someone is interested, for instance, in whether a certain type of perceptual discrimination can ever be made reliably, and how much and what type of training would be required to achieve this, then one should in the future be able to turn to a body of knowledge of documented expert performance. Our vision is that the study of expert performance will become a science of learning and of the human adaptations that are possible in response to specialized extended training. At the same time that our understanding of the real constraints on acquiring high levels of performance in any domain becomes clearer, and the similarities of those constraints across many different domains are identified, the study of the acquisition of expert performance will offer a microcosm for how various types of training can improve human performance and provide insights into the potential for human achievement.

The study of expert performance is not concerned only with the ultimate limits of performance, but also with earlier stages of development through which every future performer needs to pass. There is now research emerging on how future expert performers will acquire initial and intermediate levels of performance. Attaining these intermediate levels may be an appropriate goal for people in general and for systems of general education (e.g., recreational athletes, patrons of the arts). However, knowing how to achieve certain goals is no guarantee that people will be successful, as we know from studies of dieting and exercise. On the other hand, when the goal is truly elite achievement, the study of expert performance offers a unique source of data that is likely to help us understand the necessary factors for success, including the social and motivational factors that push and pull people to persevere in the requisite daunting regimes of training.

References


Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information


