THINKING & REASONING, 2008, 14 (4), 395–408

Psychology Press
Taylor & Francis Group

Expert and "novice" problem solving strategies in chess: Sixty years of citing de Groot (1946)

Merim Bilalić and Peter McLeod
Oxford University, UK

Fernand Gobet
Brunel University, London, UK

In a famous study of expert problem solving, de Groot (1946/1978) examined how chess players found the best move. He reported that there was little difference in the way that the best players (Grand Masters) and very good players (Candidate Masters) searched the board. Although this result has been regularly cited in studies of expertise, it is frequently misquoted. It is often claimed that de Groot found no difference in the way that experts and *novices* investigate a problem. Comparison of expert and novice chess players on de Groot's problem shows that there are clear differences in their search patterns. We discuss the troublesome theoretical and practical consequences of incorrectly reporting de Groot's findings.

Keywords: Chess; Experts vs novices; Expertise; Problem solving; Scientific citation; Search; Thinking.

Understanding the changes that take place as expertise develops is a central topic in cognitive psychology. Apart from a natural interest in what makes an expert, there is also the belief that this would help to achieve the practical goal of creating better training programmes. This is particularly true of understanding how problem-solving strategies develop as these are thought to play a central role in the acquisition of expertise (Anderson, 1993;

Correspondence should be address to Merim Bilalić, University of Tübingen, Experimental MRI, Department of Neuroradiology, Hoppe-Seyler-Str. 3, 72076 Tübingen, Germany. E-mail: Merim.Bilalic@med.uni-tuebingen.de

This research is based on a doctoral dissertation of the first author who was supported by Oxford University Clarendon and ORS scholarships. The preparation of this paper was supported by an ESRC postdoctoral fellowship to the first author.

http://www.psypress.com/tar DOI: 10.1080/13546780802265547

Koedinger & Anderson, 1990; Newell, 1980; Patel & Groen, 1991; Smith, 2002; Williams, Papierno & Makel, 2004).

One approach has been to study the way that people with different degrees of expertise tackle the same problem. An early example of this method, which has since acquired classic status, was de Groot's (1946/1965/1978) analysis of problem solving by chess players. He showed the position in Figure 1 to five Grand Masters, the most skilled players, and five Candidate Masters, very good players.[1] He asked them to think aloud while they tried to find the best move for White. On average, the Grand Masters found better moves and found them more quickly, but analysis of the verbal protocols suggested that the two groups differed little in the macrostructure of their search. Both investigated a similar number of moves and searched these to a similar depth. de Groot concluded that the superiority of the Grand Masters lay somewhere other than their search processes (for an

Figure 1. de Groot's position A (White to move). The winning move is 1. Bxd5 exd5 (1 . . . Nxd5 2. Nxd5 Bxd5 3. Bxe7; 1 . . . Bxd5 2. Bxf6 Bxf6 3.Nd7 with Nxf8) 2. Qf3 Qd8 (2 . . . Kg7 3. Ng4 Nxg4 [3 . . . Qd8 4. Bh6+] 4. Bxe7 Re8 5. Bc5 with 6. Qxg4) 3. Rce1 and the pin is decisive (e.g., 3 . . . .Kg7 4. Ng4 Ng8 5. Bxe7 Nxe7 6. Qf6+ Kg8 7. Nh6 with checkmate).

[1]Other players (three Masters, four Class C players, and two "lady players") also participated in the experiment but the key analysis (de Groot, 1978, pp. 317–320) focuses on the five Grand Masters and Candidate Masters.

explanation and historical review, see Charness, 1992; Gobet, de Voogt, & Retschitzki, 2004).

One of the attractions of chess for studying expertise is that there is an interval scale for measuring the skill levels of different players on the basis of their results against other players of known rating. Thus it is possible to compare the skill level of different groups of players precisely. The Elo scale has a theoretical mean of 1500 and a theoretical standard deviation of 200 (see Elo, 1978). Grand Masters usually have a rating of over 2500. They are at least 5 *SD* above the level of average chess players. Candidate Masters[2] have a rating between 2200 and 2000. They are about 3 *SD* above the level of average players. By the standards of psychological research on expertise, both groups of players tested by de Groot were experts. The Candidate Masters might be called "ordinary" experts as opposed to the Grand Masters who were "super" experts, but both groups were far superior to average players.[3]

Since the publication of the second English edition of de Groot's book in 1978, his work has been frequently cited (see Figure 2). The "no difference"

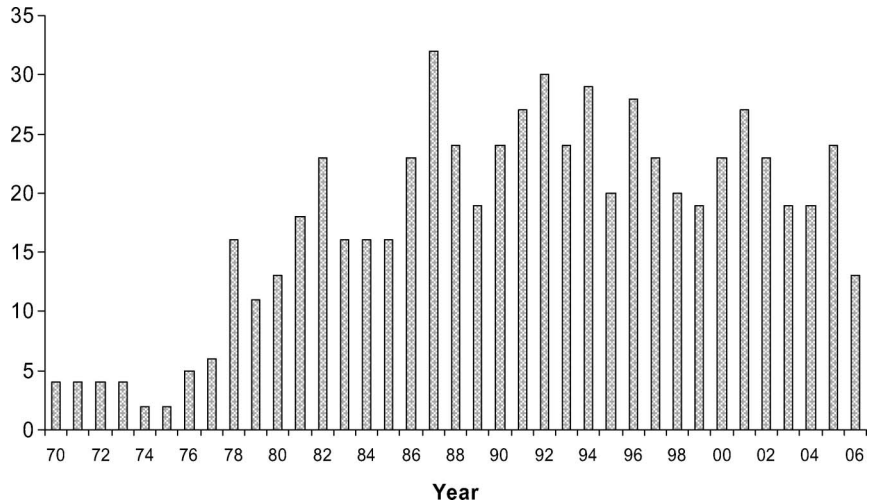

Figure 2. Frequency of citation for de Groot (1946/1965/1978) between 1970 and 2006. The data were obtained using the Social Science Citation Index and Arts & Humanities Citation Index.

[2]Players with a rating between 2000 and 2200 are sometimes called "Experts". In this paper we use the alternative title "Candidate Masters" to avoid confusion with the stronger players who are undeniably experts.

[3]The Elo scale with its associated class levels was not available when de Groot collected his data and published his thesis in 1946, but de Groot provided an explicit if rough correspondence with these levels in the English translations of his book.

search result is mentioned in many textbooks but it is often misreported. It is commonly stated that de Groot found no difference between the search processes of *experts* and *novices*. For example: "Several other investigators (de Groot, 1965; . . .) found no major strategic differences between experts and novices" (VanLehn, 1989, p. 562); "de Groot (1965) . . . found that masters . . . and novices did not differ in the number of possible moves they considered, nor how far ahead they looked" (Smyth, Morris, Levy, & Ellis, 1994, p. 358); "de Groot (1978) . . . found no reliable differences in the depth to which experts versus novices planned in advance" (Sternberg & Ben Zeev, 2001, p. 295).[4]

Another common description of de Groot's result in textbooks, less extreme than claiming that he found no difference between experts and novices but still misleading, is to report that he found no difference between *masters* and *weaker players*: "de Groot found hardly any differences between expert players and weaker players – except, of course, that the expert players chose much better moves . . . In fact, if anything, masters consider fewer moves than chess duffers" (Anderson, 2000, p. 299); ". . . de Groot (1965, 1966), who found that master players and weaker players thought about the same number of moves, considered about the same number of moves, and had a similar search patterns of moves" (Solso, MacLin, & MacLin, 2005, p. 134); "Early research by de Groot (1965) failed to uncover any evidence that expert players searched more moves, searched farther ahead, or searched faster than ordinary players" (Stillings et al., 1995, pp. 132–133). It is true that the Candidate Masters in de Groot's study were weaker players than the Grand Masters, but they were neither weak players nor ordinary, and certainly not duffers. One should, however, point out that some textbooks, such as Gilhooly (1996, p. 62), Eysenck and Keene (2000, p. 414), Green (1996, p. 334), and Parkin (2000, p. 286) accurately reproduced de Groot's finding.

Textbooks are summaries of the primary literature. Is it possible that the myth that de Groot found no differences between experts and novices stems from erroneous reporting of de Groot's result in research on expertise? Using the Social Science Citation Index and Arts & Humanities Citation Index we identified 651 journal articles from 1970 to the end of 2006 that cited de Groot's study (see Figure 2). We conducted the search for references to all three available editions of the book: 1946 – Dutch first edition; 1965 – English first edition; and 1978 – English second edition. Out of these 651 published articles we managed to examine 391 (61%). Problem-solving

[4]Curiously enough, just three pages later (p. 298) Sternberg and Ben-Zeev (2001) got it right in the context of position recall: "The assessment was further supported by de Groot's asking both the grand masters and the experts to recall a middle-game position shown to them for just a short time."

aspects of the study were mentioned in 100 (26%) of which 49 referred to details of the study. Of these, 24 (49%) used, at best, ambiguous descriptions, making the mistake repeated by the writers of textbooks cited earlier. They either stated that de Groot (a) found no differences in the problem-solving strategies between experts and novices; (b) that he used experts and novices in his study; or (c) that there were no differences between experts and weaker players (e.g., less expert, less skilled, less able, or average players). The Appendix presents the 49 articles that deal with the problem-solving aspect of the study, together with the specific misquotations.

Since de Groot did not test novices, it is possible that there *is* no difference between the search structure of experts and novices, and the misquotations are reporting a correct conclusion even if misquoting de Groot. This can be examined by comparing the performance of de Groot's Grand Masters with a group of novices who were tested on the same problem with the same instructions by Gruber (1991; see also Gobet, 1998). Gruber's novices did not have Elo ratings because they had not participated in official chess tournaments or been a member of a chess club. However, they played chess on average once a month and were not beginners. Table 1 compares the Grand Masters and Novices. Although they spent a similar amount of time on the problem, $t(27) = 1$, *ns*, $d = 0.40$, Grand Masters were, unsurprisingly, much better at finding the best move, $t(27) = 6.6$, $p < .01$, $d = 3.5$. There were clear differences in the macro-structure of search. Grand Masters reached greater maximal depths than Novices, $t(27) = 4.2$, $p < .01$, $d = 3.0$, and had more Episodes (general investigations) $t(27) = 3.4$, $p < .01$, $d = 0.98$). They did not examine more Candidate moves (different solutions) $t(27) = 0.9$, *ns*, $d = 0.44$. The small difference in this measure is probably a consequence of the position used by de Groot in which there are only a small number

TABLE 1
Average and standard deviation (in brackets) for Grand Masters (de Groot, 1946) and Novices (Gruber, 1991) on de Groot's problem

	Elo Rating	*Age (yrs)*	*Quality**	*Time (mins)*	*Max Depth**	*Candidates*	*Episodes**	*n*
Grand Masters	2658 (26)	32 (10)	4.8 (0.5)	9.4 (3.6)	7.4 (2.6)	3.6 (1.7)	6.4 (4.4)	5
Novices	–	30 (6)	1.6 (1.2)	7.8 (4.2)	3.3 (1.8)	2.9 (1.5)	3.1 (1.8)	24

Quality of solution on a scale 0 = mistake to 5 = best move (see Gobet, 1998). "Time" is time spent trying to solve the problem. "Max Depth" is the number of half moves/ply in the deepest solution searched. "Candidates" is the number of different solutions (i.e. first moves) investigated. "Episodes" is the number of general investigations. *Difference between Grand Masters and Novices significant $p < .01$.

of plausible first moves. There is also the fact that while novices are selective due to the computational cost of considering alternative moves, very strong players are highly selective in their search due to their knowledge of plausible moves.

Vicente and Brewer (1993; see also Vicente & de Groot, 1990) showed how a phenomenon from de Groot's (1946) research, that of recall of random positions, was often correctly reproduced (recall was not better in stronger players),[5] but the credit for the discovery was not. In this case, another seminal de Groot result has frequently been completely distorted. Unlike in Vicente and Brewer's case of misattribution, here we deal with incorrect reporting of a finding of potentially great importance to the understanding of expert–novice differences in problem solving. Equating the search strategies of novices and experts could have serious theoretical and practical consequences. The theories that emphasise analytical process and the role of search in expertise (e.g., Holding, 1985, 1992) would be completely discredited. Even the theories that highlight the role of pattern recognition and chunks/templates in problem solving (e.g., Chase & Simon, 1973; Gobet, 1998; Gobet & Simon, 1996; Saariluoma, 1995) would turn out to be inappropriate. Most of the mentioned theories are based not only on the seminal findings of de Groot (1946) presented here, but also in numerous others that show skill effects in search strategies (e.g., Campitelli & Gobet, 2004; Charness, 1981, 1989; Gobet, 1998; Gruber, 1991; Saariluoma, 1992).

Equally important are the practical consequences based on misreporting of the differences between experts and novices. Training programmes based on the misinterpreted results would completely abandon the emphasis on analytical abilities, one of the crucial abilities in expertise. Indeed, the field of nursing offers just such an example. S. E. Dreyfus and H. L. Dreyfus (1980) have proposed an influential theory of expertise development whose (incorrect) rejection of the role of analytical search with experts can be traced back to the misconception that there are no differences in the extent to which chess novices and masters look ahead when attempting to select a move. For example, Dreyfus and Dreyfus (1986, pp. 36–37, our emphasis) write:

> ... quality of move choice depends surprisingly little on anything but pure intuitive response.... What does a masterful chess player think about when time permits, even when an intuitively obvious move has already come spontaneously to mind? Often he uses his time to follow out sequences of moves. *Players of all levels of skill have been shown to be equally good at this.* But strong intuitive players think about other things, too.

[5]An expected twist to the story is that later research found that random positions are in fact recalled slightly but reliably better by stronger players (Gobet & Simon, 1996).

Dreyfus and Dreyfus's theory has been applied with little change to the field of nursing expertise (Benner, 1984; Benner & Tanner, 1987), and has had considerable impact on the development of curricula for nurse training. Benner (1984) clearly underplays the role of teaching analytical methods when forming expert nurses: while these methods should be taught to novices and advanced beginners (the first two of five stages of expertise in the Benner, Dreyfus, and Dreyfus model), they should not be taught in the other stages (competence state, proficiency stage, and expertise stage), where instruction should aim at developing intuitive skills. (See Gobet & Chassy, 2008, for a detailed discussion.)

Expert and novice chess players *do* differ in the macro-structure of search when trying to find the solution to a problem. The numerous textbooks and papers that cite de Groot as support for the idea that the search structure of expert and novices is similar are not only misquoting de Groot, they are putting forward an incorrect and dangerous claim. There may not be much difference between the search patterns of super experts and ordinary experts in chess positions as de Groot (1946) showed (see also Charness, 1981, 1989, but see Campitelli & Gobet, 2004); there are, however, clear differences between the search patterns of experts and novices.

Manuscript received 20 May 2008
Revised manuscript received 11 June 2008
First published online 18 August 2008

REFERENCES

Anderson, J. R. (1993). Problem solving and learning. *American Psychologist*, *48*, 35–44.

Anderson, J. R. (2000). *Cognitive psychology and its implications* (5th ed.). New York: W. H. Freeman & Co.

Benner, P. (1984). *From novice to expert: Excellence and power in clinical nursing practice*. Menlo Park, CA: Addison-Wesley.

Benner, P., & Tanner, C. (1987). Clinical judgement: How expert nurses use intuition. *American Journal of Nursing*, January, 23–31.

Campitelli, G., & Gobet, F. (2004). Adaptive expert decision making: Skilled chessplayers search more and deeper. *Journal of the International Computer Games Association*, *27*, 209–216.

Charness, N. (1981). Search in chess: Age and skill differences. *Journal of Experimental Psychology: Human Perception and Performance*, *2*, 467–476.

Charness, N. (1989). Expertise in chess and bridge. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Charness, N. (1992). The impact of chess research on cognitive science. *Psychological Research*, *54*, 4–9.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55–81.

de Groot, A. D. (1946). *Het denken van den schaker*. Amsterdam: Noord Hollandsche.

de Groot, A. D. (1965). *Thought and choice in chess*. The Hague: Mouton Publishers.

de Groot, A. D. (1978). *Thought and choice in chess* (2nd ed.). The Hague: Mouton Publishers.

Dreyfus, S. E., & Dreyfus, H. L. (1980). *A five-stage model of the mental activities involved in directed skill acquisition* (ORC 80-2). Berkeley, CA: University of California, Operations Research Center.

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

Elo, A. E. (1978). *The rating of chess players, past and present*. New York: Arco.

Eysenck, M. W., & Keane, M. T. (2000). *Cognitive psychology: A student's handbook* (4th ed.). Hove, UK: Psychology Press.

Gilhooly, K. J. (1996). *Thinking: Directed, undirected and creative* (3rd ed.). London: Academic Press.

Gobet, F. (1998). Chess players' thinking revisited. *Swiss Journal of Psychology*, *57*, 18–32.

Gobet, F., & Chassy, P. (2008). Two theories of expert intuition in nursing: A discussion paper. *International Journal of Nursing Studies*, *45*, 129–139.

Gobet, F., de Voogt, A., & Retschitzki, J. (2004). *Moves in mind*. Hove, UK: Psychology Press.

Gobet, F., & Simon, H. A. (1996). Recall of rapidly presented random chess positions is a function of skill. *Psychonomic Bulletin & Review*, *3*, 159–163.

Green, D. W. (Ed.). (1996). *Cognitive science: An introduction*. Oxford, UK: Blackwell.

Gruber, H. (1991). *Qualitative Aspekte von Expertise im Schach* [Qualitative aspects of expertise in chess]. Aachen: Feenschach.

Holding, D. H. (1985). *The psychology of chess skill*. Hillsdale, NJ: Erlbaum.

Holding, D. H. (1992). Theories of chess skill. *Psychological Research*, *54*, 10–16.

Koedinger, K. R., & Anderson, J. R. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science*, *14*, 511–550.

Newell, A. (1980). Reasoning, problem-solving and decision processes: The problem space as a fundamental category. In R. Nickerson (Ed.), *Attention and performance*, VIII (pp. 693–718). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Parkin, A. (2000). *Essential cognitive psychology*. Hove, UK: Psychology Press.

Patel, V. L., & Groen, G. J. (1991). The general and specific nature of medical expertise: A critical look. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 93–125). Cambridge, UK: Cambridge University Press.

Saariluoma, P. (1992). Error in chess: The apperception-restructuring view. *Psychological Research*, *54*, 17–26.

Saariluoma, P. (1995). *Chess player's thinking*. London: Routledge.

Smith, G. F. (2002). Thinking skills: The question of generality. *Journal of Curriculum Studies*, *34*, 659–678.

Smyth M. M., Morris P.E., Levy P., & Ellis A. W. (1994). *Cognition in action*. Hove, UK: Psychology Press.

Solso, R. L., MacLin, M. K., & MacLin, O. H. (2005). *Cognitive psychology* (7th ed.). Boston: Allyn & Bacon.

Sternberg, R. J., & Ben-Zeev, T. (2001). *Complex cognition: The psychology of human thought*. Oxford, UK: Oxford University Press.

Stillings, N. A., Weisler, S. E., Chase, C. H., Feinstein, M. H., Gareld, J. L., & Rissland, E. L. (1995), *Cognitive science: An introduction* (2nd ed.). Cambridge, MA: MIT Press.

VanLehn, K. (1989). Problem solving and cognitive skill acquisition. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 527–579). Cambridge, MA: MIT Press.

Vicente, K. J., & Brewer, W. F. (1993). Reconstructive remembering of the scientific literature. *Cognition*, *46*, 101–128.

Vicente, K. J., & de Groot, A. D. (1990). The memory recall paradigm: Straightening out the historical record. *American Psychologist*, *45*, 285–287.

Williams, M. W., Papierno, B. P., & Makel, C. M. (2004). Thinking like a scientist about real-world problems: The Cornell Institute For Research On Children science education program. *Journal of Applied Developmental Psychology*, *25*, 107–126.

APPENDIX

List of 49 articles citing de Groot's (1946/1965/1978) result on problem solving with quotations from the articles that describe the study inaccurately.

TABLE A1

Author/Year/Journal	*Vol.*	*Page*	*Quote*
(A) Studies inaccurately reporting the differences between super experts and ordinary experts as the difference between experts and novices (or unskilled/inexperienced players)			
Gonzalez et al. (2003). *Cognitive Science*	*27*	597	Experimental data from chess studies shows that experts search very selectively using recognition cues to guide their attention and achieving greater computational efficiency. Novices, however, must engage in a more thorough search to determine the principles that are applicable to the problem situation (Chase & Simon, 1973; de Groot, 1978; Simon & Gobet, 1996).
Mitchell & Dacin (1996). *Journal of Consumer Research*	*23*	220	Given that most researchers initially believed that expertise stemmed from the use of more efficient and effective search strategies, most of the early studies in this area focused on identifying differences in search strategies between domain experts and novices. Studies of chess experts, however, revealed few identifiable differences in search strategies (e.g., depth and breadth of search; de Groot 1978).
Halpern & Bower (1982). *American Journal of Psychology*	*95*	32	In his studies of chess players, de Groot (1965) found no differences between chess masters and novices in the basic reasoning associated with playing, such as the number of moves considered per play, or the depth of search for moves. However, he did find difference between the better and poorer player in a short-term memory task.
Van der Maas & Wagenmakers (2005). *American Journal of Psychology*	*3*	8	Using spoken (i.e., think-aloud) protocols during Choose-a-Move tasks, de Groot argued that skill differences are instead associated with performance in recall and recognition of standard chess positions. Skilled players presumably calculate as many moves as unskilled players, but the recognition of familiar chess patterns that drives the move selection process allows skilled players to exclude bad moves and focus their efforts only on promising continuations.

(continued)

TABLE A1
(Continued)

Author/Year/Journal	*Vol.*	*Page*	*Quote*
Bainbridge (1977). *Travail Humain*	*40*	175	For example, de Groot (1965) found that experienced chess players explore the same number of possible moves as inexperienced players, the difference is that all the moves explored by experienced players are potentionally good ones.
(B) Studies inaccurately reporting that de Groot (1946) used experts and novices			
Arocha et al. (2005). *Journal of Biomedical Informatics*	*38*	163	As is typical of experts, the network developed from the physician's explanation is fully coherent in that all nodes are connected. This contrasts with situations where a case explanation is broken into components, which is typical for novices. Similar results have been reported in other domains (Charness, 1989; Chase & Simon, 1973; de Groot, 1965).
Summers et al. (2004). *Journal of Occupational and Organizational Psychology*	*77*	293	Chess experts, for example, differ from novices both in how they encode information on positions and in their evaluation of potential future moves (e.g. Chase & Simon, 1973a, 1973b; DeGroot, 1965, 1966; Holding & Reynolds, 1982; Simon and Gilmartin, 1972)
Thompson et al. (2003). *Theory into Practice*	*42*	134	Key research identifying differences between novices and experts has been summarized by Glaser (1992) and Bransford et al. (2000). Very briefly, experts' proficiency is highly specific, and allows them to (a) discern meaningful patterns in information they encounter (e.g., the classic chess master study by deGroot, 1965) ...
Ben-Zeev & Star (2001). *Cognition and Instruction*	*19*	257	The literature on expertise shows that novices tend to focus their attention on salient surface-structural attributes of problems (attributes that are irrelevant to the solution of a problem), whereas experts perceive underlying deep-structural principles (e.g., de Groot, 1965).
Jegede & Taplin (2000). *Educational Research*	*42*	286	Early studies of expertise include those of de Groot (1965) and Chase and Simon (1973): what research tells us about the novice-expert phenomenon is that an expert differs from a novice in three ways: level of tacit knowledge, efficiency in solving problems and the application of insight in creative problem-solving.
Fletcher (1999). *Computers in Human Behavior*	*15*	394	One quite robust finding from empirical studies contrasting the approaches of expert and novice problem solvers has been that experts tend to

(continued)

TABLE A1
(Continued)

Author/Year/Journal	*Vol.*	*Page*	*Quote*
			spend much more time than novices in representing the current state of the environment or problem (Chase & Simon, 1973; deGroot, 1965; Lesgold, 1988).
Plsek (1999). *Annals of Internal Medicine*	*131*	141	Research indicates that experts in a given area do more effective thinking primarily because they have better knowledge of heuristic principles than novices do (de Groot, 1965).
Agarwal et al. (1996). *International Journal of Human-Computer Studies*	*45*	643	The influence of experience on problem solving is evident from research in expert and novice problem solving in various technical domains (de Groot, 1965; Chase & Simon, 1973; Egan & Schwartz, 1979; Larkin, McDermott, Simon & Simon, 1980). The studies indicate that experts possess chunks representing functional units in their domains, while novices do not.
Royalty (1995). *Psychological Reports*	*156*	478	The role of content-specific knowledge in critical thinking has been addressed in studies that have compared the performance of experts and novices in a number of domains, including political science (Voss, Tyler, & Yengo, 1983), chess (Chase & Simon, 1973; de Groot, 1965) ...
Stanislaw et al. (1994). *International Journal of Human-Computer Studies*	*41*	351–2	Studies of skilled performance in a variety of domains often distinguish between the cognitive processes used by experts and novices (e.g, de Groot, 1965; Ericsson & Smith, 1991) ...
Sternberg (1981). *American Psychologist*	*36*	1185	This approach, which might be referred to as a cognitive-contents approach, seeks to compare the performances of experts and novices in complex tasks such as the solution of physics problems (Chi, Feltovich, & Glaser, 1981; Chi, Glaser & Rees, in press; Larkin, McDermott, Sinon, & Simon, 1980a, 1980b), the selection of moves and strategies in chess and other games (Chase & Simon, 1978; DeGroot, 1965; Reitman, 1976), and the acquisition of domain-related ...
(C) Studies ambiguously reporting the differences between super expert and ordinary experts as the difference between experts and lesser players			
Billman & Shaman (1990). *American Journal of Psychology*	*103*	146	Classic studies by Chase and Simon (1973) showed that one dramatic difference between the more and less skilled chess players was the ability to chunk game patterns into meaningful, memorable units. Furthermore, DeGroot (1965)

(continued)

TABLE A1
(Continued)

Author/Year/Journal	*Vol.*	*Page*	*Quote*
			found that experts do not search a larger number of alternative moves.
Wolff et al. (1984). *Journal of Psychology*	*118*	7	Research on skill in the game of chess has indicated that expert and average players are similar in the way they reason about future moves sequences (de Groot, 1965) but differ in the way they perceive pieces groupings (Charness, 1976; Chase & Simon, 1973; de Groot, 1965; Frey & Adesman, 1976; Goldin, 1978).
Carter (1990). *College Composition and Communication*	*3*	124	There have been many expert-novice studies, but the seminal study, performed by deGroot on chess masters and less experienced players, best illustrates what has been learned. DeGroot began his research with the assumption that masters were better players because they could think of more possible moves and could think more moves ahead of the lesser chess players. However, de Groot discovered that neither of these hypotheses was correct: there was actually little quantitative differences between the masters and others.
Elstein et al. (1990). *Evaluation & the Health Professions*	*3*	6	The work of de Groot (1965) was especially influential; he provided models for us in two ways: the use of thinking aloud to study chess masters planning their next move and explorations of the difference between masters and weaker players.
Holding & Reynolds (1982). *Memory & Cognition*	*3*	237	Information processing models of chess skill have vainly taken their departure from de Groot's (1965) finding that players of different strengths were essentially alike in the number of moves they considered, in the depth of their search for move sequences and in other, similar measures obtained from spoken protocols.
Frey & Adesman (1976). *Memory & Cognition*	*3*	541	From an analysis of verbal protocols, de Groot (1965) established that chess Masters and less able players use similar thought processes in analyzing a complex chess position. They consider a similar number of moves (about 35), calculate to similar depths (about 7 plies), make the same number of fresh starts (about 7), and analyze a similar number of moves per minute (about 3). The major difference Groot noted was that the masters invariably analyzed stronger

(continued)

TABLE A1
(Continued)

Author/Year/Journal	*Vol.*	*Page*	*Quote*
			moves than the weaker players.
Lawson (2004). *Design Studies*	*3*	448	de Groot's work showed that a key distinguishing factor between the chess master and the less expert player was as much in perception as in action. Chess masters, he found, rarely analysed a board situation, rather they recognised it. He showed that chess masters could remember mid-game board situations much more reliably than novices. However, their comparative expertise vanished when asked to remember randomly positioned pieces that did not relate to game situations. Taken together these results suggest something we are familiar with in design, the use of known precedents that have been studied and about which the expert has schemata. These precedents linked problem to solution and such chess masters could articulate this link. Thus, the schema for the situation also includes one or more known gambits for solving it.
Schultetus & Charness (1999). *American Journal of Psychology*	*3*	555	One the first attempts to explore the role of each process in chess was a study by de Groot (1965) in which he analyzed recall performance and verbal protocols of expert and club-level chess players. Although he found no skill-related differences in depth of search, he reported that skilled players were capable of reproducing game positions with accuracy.
Gilhooly (1990). *Applied Cognitive Psychology*	*4*	262–263	de Groot found that while expert and less expert chess players did not differ in the amount of mental search in which they engaged before choosing a move, they did differ in memory for briefly presented chess positions.

TABLE A2
Articles correctly reproducing the finding

Name/Year/Journal	*Vol.*	*Page*
Shepherd et al. (2006). *Psychology & Marketing*	*23*	115
Fazey et al. (2006). *Ecology and Society*	*10*	2
Burns (2004). *Psychological Science*	*15*	443
Campitelli & Gobet (2004). *ICGA journal*	*27*	209
Gobet & Clarkson (2004). *Memory*	*12*	732
Sweller (2004). *Instructional Science*	*32*	11
Gobet (1998). *Swiss Journal of Psychology*	*57*	19
Sinclair (1998). *Advances in Case-Based Reasoning*	*1488*	127
Berliner & McConnell (1996). *Artificial Intelligence*	*86*	99
Ericsson & Lehmann (1996). *Annual Review of Psychology*	*47*	278
Gobet & Simon (1996). *Psychological Science*	*7*	52
Lindgaard (1995). *Interacting with Computers*	*7*	257
Burns (1994). *Behavioral and Brain Sciences*	*17*	535
Reynolds (1991). *Bulletin of the Psychonomic Society*	*29*	55
Klein & Peio (1989). *American Journal of Psychology*	*102*	322
Bransford et al. (1986). *American Psychologist*	*41*	1079
Holding & Pfau (1985). *American Journal of Psychology*	*98*	271
Boshuizen & Claessen (1982). *Medical Education*	*16*	84
Reynolds (1982). *American Journal of Psychology*	*95*	384
Charness (1981). *Journals of Gerontology*	*7*	467
Charness (1981). *JEP: HP&P*	*36*	618
Erickson & Jones (1978). *Annual Review of Psychology*	*29*	72
Simon & Chase (1973). *American Scientist*	*62*	396
Chase & Simon (1973). *Cognitive Psychology*	*4*	55