

## ADAPTIVE EXPERT DECISION MAKING: SKILLED CHESS PLAYERS SEARCH MORE AND DEEPER

*Guillermo Campitelli*<sup>1</sup> *Fernand Gobet*<sup>1</sup>

Uxbridge, United Kingdom

### ABSTRACT

Previous research has suggested that depth of search in chess does not increase much as a function of skill. We submitted players to a problem-solving task with complex positions. We found a strong skill effect in depth of search, rate of search, and number of nodes generated. At the level of strong masters, the absolute values of these variables were much higher than in previous studies (sometimes 10 times higher). Supplementary data on memory, practice, reaction times, and time-constrained decision making (a maximum decision time of 10 seconds) indicated that players' behaviour was consistent with the behaviour of players previously studied in the literature and with predictions of theories based on pattern recognition. Beyond adding support to the hypothesis that both the ability to search and pattern recognition are relevant aspects of expert thinking, these results are important in showing that previous research has vastly underestimated experts' search potential. We conclude that long-term memory knowledge allows both extensive search and rapid evaluation when making decisions under time pressure. Players adaptively use either problem-solving method depending on the demands of the task.

### 1. INTRODUCTION

The formal description of problem solving as a search through a problem space (Newell and Simon, 1972) raises the question as to how much of this space must be searched before reaching a decision. While computers must search problem spaces extensively to solve problems (Russell and Norvig, 1995), as is apparent for example in game playing (e.g., Berliner, 1984; van den Herik and Iida, 1998), humans carry out searches orders of magnitude smaller, in part due to their limited computational resources. In his theory of bounded rationality, Simon (1947, 1955) argued that humans choose “good enough” solutions with little search because the heuristics they use make their search highly selective.

Independently of Simon, De Groot (1946/1978) addressed the question of bounded rationality in his seminal study on chessplayers' thinking and decision making. Participants were presented with a chess position and were required to think aloud while pondering their next move. The surprising finding was that there were no differences in the macro-structure of search between grandmasters (including world champions) and candidate masters, although the former prevailed over the latter in terms of the quality of the chosen moves. That is, number of moves generated, mean depth of search, and maximal depth of search — among other variables — remained roughly constant across the different skill levels. By contrast, De Groot found large quantitative differences between strong masters and weaker players in a memory task.

These results led to the development of pattern-recognition theories of expertise (e.g., chunking theory: Chase and Simon, 1973; template theory: Gobet and Simon, 1996a), according to which the storage of a large number of perceptual chunks (familiar configurations) in LTM enables fast recognition of key configurations in the problem at hand and thus selective search. While mostly corroborated by data on perception and memory (Gobet, de Voogt, and Retschitzki, 2004), these theories have also received support from investigations of problem solving. Gobet and Simon (1996b) found that world champion Kasparov did not decrease his performance substantially when playing simultaneous games, even though the reflection time was considerably diminished. Klein, Wolf, Militello, and Zsombok (1995) showed that, in a decision-making task, the first move that players generated was usually “good enough”. Burns (2004) found that performance in blitz (about 5 s per

---

<sup>1</sup> Centre for Cognition and Neuroimaging, Brunel University, Uxbridge, Middlesex, UB8 3PH, United Kingdom.  
Email: {Fernand.Gobet,guillermo.campitelli}@brunel.ac.uk.

move, on average) shares 81% of the variance with the ratings based on standard chess (about 180 s per move, on average). He also suggested that, as players' skill increases, proportionally more of that skill is based on fast processes such as pattern recognition than on slow processes such as search. Finally, Charness, Reingold, Pomplun and Stampe (2001) showed a skill effect (candidate masters better than intermediate players) in a simple problem-solving task, both in accuracy and speed of response.

Holding (1985) criticized the pattern-recognition approach, suggesting that search and evaluation, rather than pattern recognition, are the key processes of expert performance. This opposition was misguided, as pattern recognition and search are not mutually exclusive in pattern-recognition theories, but complementary (Chase and Simon, 1973; Gobet and Simon, 1998). Search is enabled by the recursive application of pattern-recognition processes in the mind's eye. Once a pattern is recognized and a move generated, the image in the mind's eye incorporates this change, and pattern recognition takes place over the new image. This idea is implemented in SEARCH (Gobet, 1997), a formal model that combines pattern recognition, search, and mental imagery. Simulations with SEARCH showed that the interaction of recognition and search leads to an increase in depth of search as a function of skill. Indeed, this prediction is similar to Holding's (1985), who argued that search should increase almost linearly with expertise. Both predictions can be contrasted with Charness's (1981), who proposed that depth of search reaches a plateau after candidate master level.

On the whole, replications of De Groot's experiment have supported Holding's prediction only weakly. Charness (1981) submitted players ranging from 1284 to 2004 Elo, to four different positions. He found that the quality of moves was a function of skill and that there were slight but statistically significant skill differences in depth of search. There was also a skill effect in total moves analyzed, number of episodes, and number of branches. However, no effect was found in rate of search. Gobet (1998) reused De Groot's position A with players ranging from 1600 to 2450 Elo. The quality of moves and mean depth varied across levels of expertise; however, the differences were rather low and no effect was found in maximal depth of search. He also found that masters searched faster than candidate masters<sup>2</sup>. Saariluoma (1990) examined players from 1900 to 2500 Elo and found that international masters sometimes searched less deeply than master players. Finally, Chabris and Hearst (2003) found that top-level grandmasters made fewer blunders (5.02 per 1,000 moves) in standard chess than in rapid chess (6.85). In rapid chess, about 30 s are allotted per move, on average.

Thus, although both pattern-recognition and search theories predict a positive skill effect with search variables, previous studies found no skill effect, only a small effect, or even a negative effect. The lack of clear skill differences in absolute or relative terms is even more surprising given that some grandmasters report searching at great depths in competitive games (e.g., Shirov, 1997). A possible explanation for this paradox resides in variations in task difficulty (Saariluoma, 1995) and in the fact that researchers have used positions whose solution does not require extensive search. For example, based on the solution provided by De Groot (1978), his position A can be solved by visiting only 17 nodes<sup>3</sup>, with a maximal depth of 9 ply.

Here, we present new data showing that, consistent with Holding's claim (1985), chessmasters' search potential had been underestimated for almost fifty years. We followed De Groot's (1978) methodology but selected complex positions whose unique solution required extensive search. We seconded these results with data on fast decision making, memory, and practice. This allowed us to ascertain that our players did not differ from players previously studied and that their skill (partly) depends on pattern recognition. Given that we expected large differences between individuals of various skill levels in our study, we used only four participants (using a small number of participants has a long tradition in expertise research; e.g., Chase and Simon, 1973; Wagner and Scurrah, 1971).

---

<sup>2</sup> Elo (1978) developed the rating scale that is now used by the World Chess Federation (FIDE). The scale has a normal distribution and a standard deviation of 200 points. The best player of the world has around 2800 points and the weakest 1200. In the psychology literature, players between 1600 and 1800 are called Class B, between 1800 and 2000 Class A, between 2000 and 2200 experts, and players with more than 2200 are considered masters. FIDE awards players with titles for their performances in specific tournaments. As an approximation, players above 2300, 2400 and 2500 are called FIDE masters, international masters, and international grandmasters, respectively. In this article, we use "candidate master" instead of "expert" to avoid confusion with the general meaning of the term "expert".

<sup>3</sup> A node is one position in the search tree (either that produced in a think-aloud protocol or that generated by analysis).

## 2. RESEARCH METHOD

Below we discuss our research method by describing the participants (2.1), the materials (2.2), and the procedure (2.3).

### 2.1 Participants

One international grandmaster (GM) with 2550 Elo points (age 21), one international master (IM, 2500, age 22), one candidate master (CM, 2100, age 19), and one class B player (CB, 1750, age 19) participated in our study.

### 2.2 Materials

The three positions used (see Figure 1) were complex, requiring considerable look-ahead search and evaluation. Position A was De Groot's (1978) position C. He presented a possible solution requiring 52 nodes, with a maximal depth of 23 ply<sup>4</sup>. The second position required at least 60 nodes, with a maximal depth of 25 ply (Roycroft, 1972, #188). The third position required at least 70 nodes, with a maximal depth of 35 ply, to be solved (Nunn, 1999, #3). We did not use De Groot's position A, as it is now widely available in the chess literature, making it likely that our players had already seen it. The main criterion for choosing the positions was that their solution required deep search. This was essential for our study because the lack of a strong relationship between chess skill and depth of search in previous studies might have been caused by the use of positions that did not require deep search.

A. Black moves.

B. White moves and wins.

C. White moves and wins.

**Figure 1.** Positions used in the thinking-aloud task (Experiment 1). Position (a) was used by De Groot (1978), position (b) is a problem created by Kasparyan (Roycroft, 1972, #188), and position (c) was obtained from Nunn (1999).

For the three positions, the main lines are as follows. Position A: De Groot proposes that Black can free himself by 1. ...e5 2. Bxc8 Raxc8! 3. dxe5 Qe4! 4. Rhe1 Qxe2 5. Rxe2 Ng4! 6. exd6 Rxf3, with both 7. dxc7 and 7. d7 failing to give an advantage to White. However, FRITZ finds that 7. Na7! gives White an edge (Merim Bilalić, personal communication). Position B: The main line is 1. Bg5 b3 2. Rd2+ Ka1 3. f7 Rxf5 4. f8=Q Rg1+ 5. Rd1 Rg2 6. Qa3+ Ra2 7. Rd2 Rxa3 8. Rb2 Ra2 9. Rb1 mate. Position C: The game Polugaevsky-Torre (London, 1984) continued by 1. Bxh7+ Kxh7 2. Qh5+ Kg8 3. Rg3 g6 4. Rxf6+ fxg6 5. Qxg6+ Kh8 6. Qh6+ Kg8 7. Qxe6+ Kh8 8. Qh6+ Kg8 9. Qg6+ Kh8 10. Qh5+ Kg8 11. Bh6 Bf8 12. Qg6+ Kh8 13. Bxf8 Rxf8 14. Qh6+ Kg8 15. Ra3 1-0.

### 2.3 Procedure

The participants were required to put themselves in a tournament situation as if they were to move. Thirty minutes were allowed as maximum reflection time for each position. When the players had reached a decision, they had to play the chosen move. Following De Groot's instructions, players were required to say aloud everything they were thinking of (e.g., moves, plans, evaluations, and even irrelevant thoughts). Positions B

<sup>4</sup> A ply is the level of depth in a sequence of moves. In chess it corresponds to half a move. A move consists of two ply, one White and one Black.

and C had a unique solution requiring deep and wide search (participants were told that they had to find it). Concurrent think-aloud protocols were tape recorded.

### 3. RESULTS

Once the protocols were transcribed, we calculated the values of three variables: depth of search, number of nodes generated, and rate of search. As in previous experiments, depth of search was defined as the number of ply generated from the problem position until the end of one variation. This measure was further refined as average depth of search and maximal depth of search. Rate of search was defined as the number of nodes generated divided by the reflection time. Table 1 displays the average results across the three positions for each player.

	GM		IM		CM		CB	
	m	sd	m	sd	m	sd	m	sd
Quality of moves	4.3	1.1	4.7	0.6	3	1.7	3	1.4
Total time (min)	30	0	30	0	22.5	3.4	28.9	1.5
#nodes generated	370	101	405	44	152.7	30	88	2.8
Rate of search (nodes/min)	12.3	4.1	13.5	1.7	6.8	1.2	3.2	0.2
Max. depth (in ply)	25	12.2	23.7	5.1	17.7	4.9	10.5	2.2
Mean depth (in ply)	13.8	7.4	10.5	5.0	7.4	2.8	2.8	0.2

**Table 1:** Means and standard deviations of variables obtained from thinking aloud protocols. For each player, the data are averaged across all positions.

#### 3.1 Depth of search

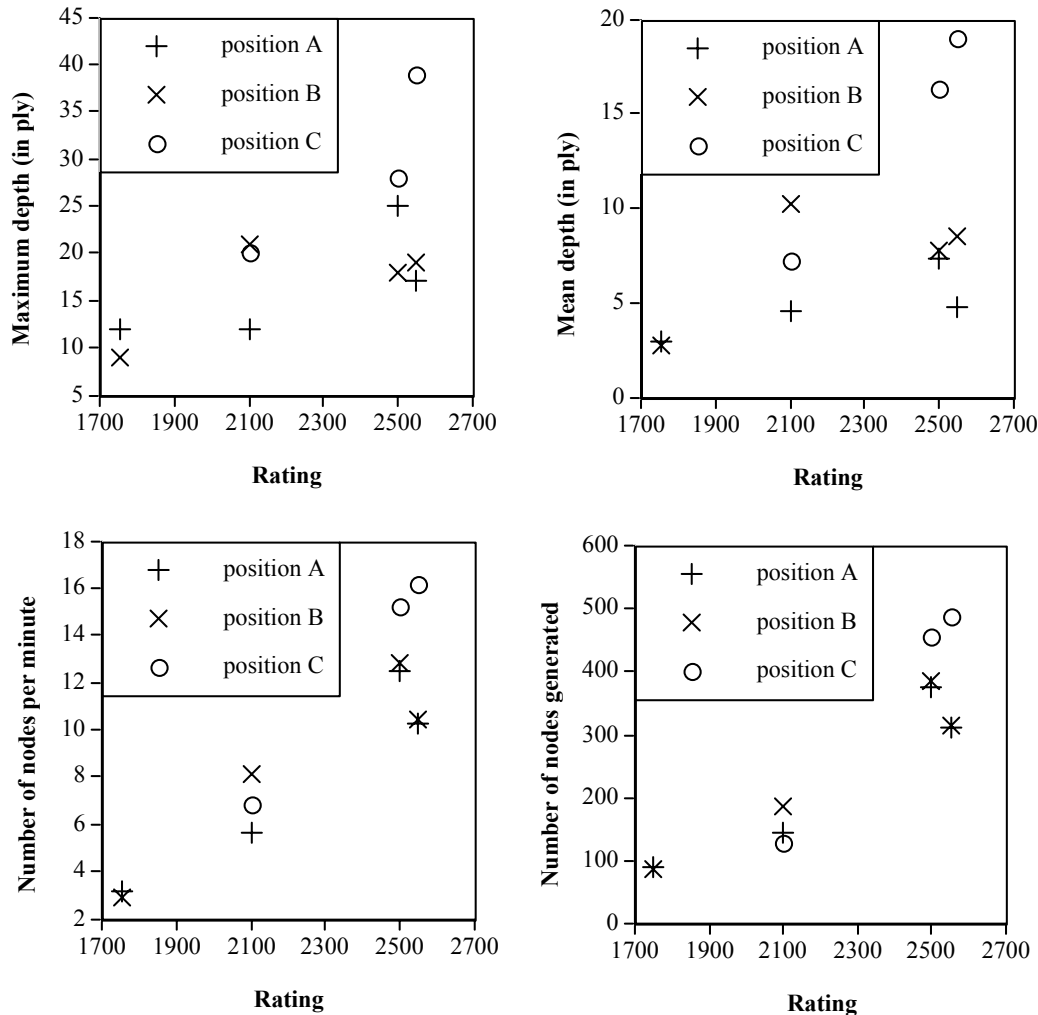
In all measures of depth there was a strong linear relation with chess rating (see Figure 2, top left and top right). Regression equations were calculated for mean and maximal depth of search (for each player, results were averaged across the three positions), using chess rating as a predictor. With mean depth, the regression line was: mean depth =  $-18.3 + 0.012 * \text{Elo}$  ( $r(4) = .97$ ,  $p < 0.02$ , one-tailed). Mean depth of search thus increases 2.4 ply for each standard deviation of chess skill (200 Elo points), which is larger than the 0.5 and 0.6 ply obtained by Charness (1981) and Gobet (1998), respectively. With maximal depth of search, the equation was: maximal depth =  $-19.9 + 0.018 * \text{Elo}$  ( $r(4) = .99$ ,  $p < .002$ , one-tailed). The increase per standard deviation was 3.6 ply, which is higher than the 1.5 ply found in Charness (1981); Gobet (1998) did not find skill differences in this variable.

Remarkably, the absolute values of the stronger players were more than twice those of players of the same level in previous studies. The mean depth of the top players in De Groot (1978), Saariluoma (1990) and Gobet (1998) was 5.4, 3.6 and 5 ply respectively, whereas in our study, GM and IM's mean depth were 13.8 and 10.5 ply, respectively. Similarly, maximal depth for the top category was 6.8 and 9.1 ply in De Groot's and Gobet's studies, respectively (no maximal depth was calculated in Saariluoma, 1990), whereas GM and IM's maximal depth was 25 and 23.7 ply, respectively.

#### 3.2 Rate of search

There was a strong skill effect in rate of search (see Figure 2, bottom left). GM generated 12.3 nodes per minute; IM 13.5; CM 6.8; and CB 3.2. Predicting rate of search from chess rating yields the following equation: Rate of search =  $-19.2 + \text{Elo} * 0.013$  ( $r(4) = .98$ ,  $p < .01$ , one-tailed), which corresponds to an increase of 2.6 nodes per minute per Elo standard deviation. The absolute values obtained at the stronger level for rate of search were three times higher than those found in previous studies (De Groot's (1978) and Gobet's (1998) top players searched at a rate of 3.6 and 4.8 nodes per minute, respectively). However, Wagner and

Scurrah (1971) reported data of a candidate master (incidentally, the authors wrongly indicated that the player was class C) searching at an average of 10 nodes per minute in five different positions, which is a higher value than that of our candidate master.



**Figure 2:** Top left: Maximal depth (in ply). Top right: mean depth (in ply). Bottom left: rate of search (nodes per minute). Bottom right: number of nodes generated, as a function of chess rating (Experiment 1). The data of the three positions are presented (because of a technical problem, CB's data on position C were lost). CBs data points correspond to 1750 in the x-axis, CM's to 2100, IM's to 2500, and GM's to 2550.

### 3.3 Number of nodes generated

A statistically significant linear relation between number of nodes generated and chess skill was found (GM 370 nodes; IM 405; CM 153; and CB 88; see Figure 2, bottom right). Using Elo to predict the number of nodes generated yields the following formula. Number of nodes generated =  $-645 + \text{Elo} \cdot 0.4$  ( $r(4) = .96$ ,  $p < .02$ , one-tailed), which corresponds to an increase of 80 nodes per Elo standard deviation. At the level of strong masters, the absolute values presented here (an average of 387.5 nodes) are ten times higher than those found in De Groot's, Gobet's and Saariluoma's studies, where top players generated only 35, 58 and 23 nodes, respectively (an average of 38.6 nodes). Wagner and Scurrah's (1971) candidate master generated an average of 172 nodes; this is similar to the 153 nodes generated by our candidate master.

#### 4. QUALITATIVE FEATURES OF THE PROTOCOLS

In his influential book written for the chess community, Kotov (1971) suggested that grandmasters are better than weaker players because they know how to calculate variations and are more disciplined in their thinking. Based on this observation, he proposed a method for helping players to carry out search more efficiently. The first step is to decide which moves are worth considering (candidate moves). Once this is done, the player has to analyze each move until an evaluation of the position can be produced, without repeating the analysis of a given variation or “jumping” from one branch of the search tree to another. Thus, a new candidate move can be analyzed only when the analysis of the previous candidate move is completed. Interestingly, this is not what De Groot (1946) found: all the players he studied, including the grandmasters, tended to reinvestigate the same variation several times, either immediately or after the consideration of a different variation, a phenomenon he called “progressive deepening”. Given the discrepancy between Kotov’s advice and De Groot’s data, it is of interest to look at this aspect of the search behaviour in our sample (see also Gobet and Jansen, in press, for a further discussion of this question).

The protocols indicate that CM and CB did not generate a list of candidate moves and “jumped” from one move to another. Surprisingly, GM also jumped between moves; however, there were two differences between his search behaviour and that of the intermediate players. First, GM generated a list of candidate moves that was enlarged when a new idea, for which a new candidate move was necessary, occurred to him. Second, he was aware that he was analyzing without discipline and tried to compose himself while he was performing the task. Finally, IM showed a “textbook” behaviour: he generated a list of candidate moves which he modified very rarely, and tended to analyze systematically each move until he had reached an evaluation. In sum, there were clear differences between the masters and the intermediate players in the knowledge of how chess trainers, such as Kotov, recommend analysis of variations to be conducted and in the actual search behaviour. It may be worth noting that both IM and GM had read Kotov’s book. These results suggest that search behaviour is not restricted to progressive deepening but can be under strategic control.

#### 5. ADDITIONAL DATA

Before discussing the theoretical impact of our results, we need to make sure that our players are representative of those used in previous studies. For this purpose, we briefly report data collected with them on rapid decision making, memory, and practice. If the skill effects observed in search behaviour remain in these tasks, the hypothesis of compatibility between pattern recognition and ability to search deeply would be supported.

The fast-decision experiment imposed drastic constraints on thinking time, so that players were able to perform only limited search and presumably had to rely more on pattern recognition. Forty-nine positions of medium complexity were selected from Livshitz (1988); they required 10.4 nodes ( $sd = 4.2$ ) on average to be solved, with a depth of search of 6.4 ply ( $sd = 1.6$ ). Each position was presented on a computer screen for 5 s, followed by a 5 s black screen. Within this 10-second period, participants had to verbalize their move. A clear skill effect was found: GM correctly solved 23 problems (46.9%), IM 18 (36.9%), CM 4 (8.2%), and CB 2 (4.1%). The correlation between Elo rating and the accuracy in the task was significant ( $r(4) = .94$ ;  $p < .03$ , one-tailed).

To control for the possibility that these results were due to differences in domain-general perceptual abilities, we asked participants to carry out a simple reaction-time task. A red or a yellow circle was presented on the screen, and participants had to press the relevant button as fast and accurately as possible. No skill differences were found: GM’s mean was 339 ms, with a standard deviation of ( $\pm$ ) 70 ms; IM, 369 ms  $\pm$  66 ms; CM, 357 ms  $\pm$  61 ms; and CB, 343 ms  $\pm$  75 ms.

If the skill effect observed in problem-solving tasks can be accounted for by domain-specific LTM patterns, as proposed by Chase and Simon, then there should also be a skill effect in a memory task requiring the use of these patterns. To test this hypothesis, we submitted our participants to De Groot’s (1978) recall task, which has been previously used in numerous studies (see Saariluoma, 1995, or Gobet et al., 2004). We used three types of positions: game, random, and shape (this condition consisted in showing shapes instead of chess pieces). The positions were displayed on a computer screen for 5 s. We found a strong skill effect in the game condition (83.2%, 85%, 54%, and 47%, for GM, IM, CM, and CB, respectively), a small effect in the random condition (21.7%, 19%, 17.7%, and 15.5%) and no differences in the shape condition (15.9%, 15.5%, 15%, and 15%). These results are consistent with previous data (see Gobet and Simon, 1996c).

Since Chase and Simon (1973) hypothesized that domain-specific patterns are acquired by practice, we asked our participants to fill in a grid stating how many hours per week they spent either studying or practicing chess in each year of their chess career (see Charness, Krampe and Mayr, 1996, for a similar approach). With this information we calculated the cumulative time spent studying and playing chess. Consistent with the literature, the results showed a strong skill effect. For studying, the cumulative hours spent by GM, IM, CM and CB were: 6890, 7904, 1872, and 416 hours, respectively; for playing: 7722, 7072, 2704, and 1326 hours, respectively.

## 6. GENERAL DISCUSSION

In this paper, we submitted players of various skill levels to complex chess positions. Contrary to De Groot's (1978) results, a clear skill effect was found in the three search variables investigated. We also found a skill effect when players had to solve problems under extreme time constraints. In addition, consistent with the literature, the strong players of our study (IM and GM) performed better than the intermediate players (CM and CB) in a memory task with briefly presented chess positions and spent more time practising their skill, but did not show evidence of general skills in a reaction-time task.

While Charness (1981) and Gobet (1998) did find small skill effects in depth of search, the novelty of our study is that it shows a linear relation between depth of search and Elo rating, from class B to grandmaster. The absolute values at the level of strong masters were from two to ten times higher than those reported in the literature. These results are inconsistent with Charness's hypothesis that search reaches an asymptote when players become masters, but support Holding's (1985) hypothesis that it increases almost linearly with skill. They are also consistent with SEARCH's (Gobet, 1997, Figure 3) predictions that search on average increases as a power function of skill, although the rate of increase is much higher than that proposed by SEARCH.

These results support our hypothesis that positions requiring large trees of moves to be solved lead to substantial skill differences in search behaviour. Our results are important because they correct a widespread underestimation in the psychological literature of the search capabilities of chess professionals. An interesting question for further research is whether this result generalizes to other domains of expertise.

While able to search deep in our think-aloud experiment, our players behaved as predicted by pattern-recognition theories in the memory and fast-decision tasks. Remarkably, in the second task, GM and IM were able to solve one half and one third of the problems, respectively, although the decision had to be reached within 10 seconds. How can these results be reconciled? Pattern-recognition theories predict skill differences in search variables when the position requires deep and wide search (as we found in our study), but not when the position can be solved via pattern recognition without extensive search. In this study, we avoided the circularity that this reasoning may suggest by reporting the number of nodes and ply required to solve positions, based on published sources.

We acknowledge some weaknesses of our study. The size of our sample is not large enough to make precise quantitative estimates of the search variables under study. Although the results of the fast-decision task make us reasonably confident that IM and GM can solve problems of medium difficulty without extensive search, we did not use a relatively simple position (such as De Groot's position A) with the think-aloud task. Furthermore, with our positions B and C, we encouraged our participants to find a unique solution, which is an addition to De Groot's instructions. This may have affected search behaviour, beyond the characteristics of the positions used. While we believe that these shortcomings do not invalidate our main result — that previous literature has vastly underestimated chess experts' search abilities — we agree that they obscure the conditions requiring more exhaustive or more selective search. Further research is required to clarify these conditions.

We conclude that Holding (1985) was right in claiming that chessmasters (and presumably other experts) can sometimes carry out extensive search. However, De Groot (1978) and Chase and Simon (1973) were also correct in highlighting the importance of selective search and the role of pattern recognition. The strength of pattern-recognition theories, which propose that pattern recognition is an important part of look-ahead search, is that they provide mechanisms for both rapid and slow decision making (Saariluoma, 1995; Gobet, 1997; Gobet and Simon, 1998). In contrast, search theories (e.g., Holding, 1985), while supported by the results of our think-aloud experiment, fail to explain how skill effects remain when there is almost no opportunity for search.

## 7. ACKNOWLEDGEMENTS

We thank Merim Bilalić, Bruce Burns, Neil Charness, Peter Lane, Amanda Parker, Pertti Saariluoma, Henk Visser, and anonymous referees for their useful comments on this study.

## 8. REFERENCES

- Berliner, H. J. (1984). Search vs. knowledge: An analysis from the domain of games. In A. Elithorn and R. Banerji (Eds.), *Artificial and Human Intelligence* (pp. 105-117). New York, NY: Elsevier.
- Burns, B. D. (2004). The effects of speed on skilled chess performance. *Psychological Science* 15, 442-447.
- Chabris, C. F., and Hearst, E. S. (2003). Visualization, pattern recognition, and forward search: Effects of playing speed and sight of the position on grandmaster chess errors. *Cognitive Science*, 27, 637-648.
- Charness, N. (1981). Search in chess: Age and skill differences. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 467-476.
- Charness, N., Krampe, R. and Mayr, U. (1996). The role of practice and coaching in entrepreneurial skill domains: An international comparison of life-span chess skill acquisition. In K. A. Ericsson (Ed.) *The road to excellence*. (pp. 51-80). Mahwah, NJ: Erlbaum.
- Charness, N., Reingold, E. M., Poplun, M., and Stampe, D. M. (2001). The perceptual aspect of skilled performance in chess: Evidence from eye movements. *Memory and Cognition*, 29, 1146-1152.
- Chase, W. G., and Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press.
- De Groot, A. D. (1946/1978). *Thought and choice in chess*. The Hague: Mouton. (2nd English ed.; first Dutch edition published in 1946)
- Gobet, F. (1997). A pattern-recognition theory of search in expert problem solving. *Thinking and Reasoning*, 3, 291-313.
- Gobet, F. (1998). Chess players' thinking revisited. *Swiss Journal of Psychology*, 57, 18-32.
- Gobet, F., de Voogt, A., and Retschitzki, J. (2004). *Moves in mind: The psychology of board games*. Hove, UK: Psychology Press.
- Gobet, F. and Jansen, P. J. (in press). *Training in chess: A scientific approach*. In T. Redman (Ed.), Education and chess.
- Gobet, F., and Simon, H. A. (1996a). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology*, 31, 1-40.
- Gobet, F., and Simon, H. A. (1996b). The role of recognition processes and look-ahead search in time-constrained expert problem solving: Evidence from grand-master-level chess. *Psychological Science*, 7, 52-55.
- Gobet, F., and Simon, H. A. (1996c). Recall of rapidly presented random chess positions is a function of skill. *Psychonomic Bulletin and Review*, 2, 159-163.
- Gobet, F., and Simon, H. A. (1998). Pattern recognition makes search possible: Comments on Holding (1992). *Psychological Research*, 61, 204-208.
- Holding, D. H. (1985). *The psychology of chess skill*. Hillsdale, NJ: Lawrence Erlbaum.
- Klein, S. W., Wolf, S., Militello, L., and Zsombok, C. (1995). Characteristics of skilled option generation in chess. *Organizational Behavior and Human Decision Processes*, 62, 63-69.
- Kotov, A. (1971). *Think like a grandmaster*. London: Batsford.
- Livshitz, A. (1988). *Test your chess IQ* (Vol. 1). Oxford, UK: Pergamon Press.
- Newell, A., and Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.
- Nunn, J. (1999). *John Nunn's chess puzzle book*. London: Gambit.
- Roycroft, A. J. (1972). *Test tube chess*. Harrisburg, PA: Stackpole books.
- Russell, S., and Norvig, P. (1995). *Artificial intelligence: A modern approach*. Englewood Cliffs, NJ: Prentice Hall.
- Saariluoma, P. (1990). Apperception and restructuring in chess players' problem solving. In K. J. Gilhooly, M. T. G. Keane, R. H. Logie and G. Erdos (Eds.), *Lines of thinking* (Vol. 2, pp. 41-57). New York: John Wiley.
- Saariluoma, P. (1995). *Chess players' thinking*. London: Routledge.
- Shirov, A. (1997). *Fire on board*. London: Cadogan.
- Simon, H. A. (1947). *Administrative Behavior*. New York, NY: Macmillan.
- Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, 69, 99-118.
- van den Herik, H. J., and Iida, H. (1998). *Computers and Games* (Lecture Notes in Computer Science, Vol. 1558). Berlin: Springer-Verlag.
- Wagner, D. A., and Scurrah, M. J. (1971). Some characteristics of human problem-solving in chess. *Cognitive Psychology*, 2, 454-478.