

Running Head: EXPERTISE AND INTUITION

Expertise and Intuition: A tale of three theories

Fernand Gobet and Philippe Chassy

Centre for the Study of Expertise

Brunel University

Address correspondence to:

Fernand Gobet
Centre for the Study of Expertise
Brunel University
Uxbridge, Middlesex, UB8 3PH
United Kingdom

Phone: +44 (1895) 265484

Fax: +44 (1895) 237573

fernand.gobet@brunel.ac.uk

Abstract

Several authors have hailed intuition as one of the defining features of expertise. In particular, while disagreeing on almost anything that touches on human cognition and artificial intelligence, Hubert Dreyfus and Herbert Simon agreed on this point. However, the highly influential theories of intuition they proposed differed in major ways, especially with respect to the role given to search and as to whether intuition is holistic or analytic. Both theories suffer from empirical weaknesses. In this paper, we show how, with some additions, a recent theory of expert memory (the template theory) offers a coherent and wide-ranging explanation of intuition in expert behaviour. It is shown that the theory accounts for the key features of intuition: it explains the rapid onset of intuition and its perceptual nature, provides mechanisms for learning, incorporates processes showing how perception is linked to action and emotion, and how experts capture the entirety of a situation. In doing so, the new theory addresses the issues problematic for Dreyfus's and Simon's theories. Implications for research and practice are discussed.

Acknowledgments

We are grateful to Stuart Dreyfus, Gary Klein, Pat Langley and Richard Smith for detailed comments on a previous version of this manuscript.

Keywords

Chunking; consciousness; decision making; emotions; expertise; intuition; memory; pattern recognition; template

Expertise and Intuition: A tale of three theories

One of the hallmarks of expertise is the speed and ease with which experts can recognize the key features of a situation, a phenomenon often called intuition.¹ For example, a radiologist can diagnose a disease nearly instantaneously, and a chess grandmaster can literally ‘see’ the good move straight away. With routine problems, the decision will be correct most of the time. This phenomenon has attracted wide attention in the literature, which has been dominated by two main theories. On the one side, authors such as Hubert Dreyfus (Dreyfus, 1972; Dreyfus & Dreyfus, 1988) have argued that intuition is a signature of the holistic processing of the brain and the mind.² On the other side, authors such as Herbert Simon (Chase & Simon, 1973; Simon, 1989) have proposed that simple mechanisms, based on pattern recognition, are sufficient for explaining intuition. In spite of these differences, it is important to note that Dreyfus and Simon agreed on many aspects of intuition: its speed, its fluidity, the fact that it takes a large amount of practice for a novice to reach expert level and thus show intuitive behaviour, and the fact that perceptual processes lie at the core of intuition. This level of agreement is rather ironic—and often ignored in the literature—given that these two scholars were at the centre of a bitter dispute as to whether artificial intelligence (AI) was myth or reality, the use of symbols in human cognition, and the importance of heuristics in decision making.

¹ In line with the literature, we use *intuition* for the rapid understanding shown by individuals, typically experts, when they face a problem, and *insight* for the sudden discovery of a solution after a protracted and unsuccessful search. While this article focuses on intuition, several of the theories we discuss have been applied to explain insight as well.

² As the authors made a number of corrections, we use the 1988 edition of *Mind Over Machine* rather than the 1986 edition.

The goal of this article is to evaluate these two theories empirically and then to present a new theory of intuition that removes their limitations. We first present the empirical evidence supporting the psychological reality of the concept of intuition, then we discuss Dreyfus's and Simon's theories in detail. The identification of the strengths and weaknesses of these two theories leads to the presentation of a new theory of intuition based on the template theory of expertise (Gobet & Simon, 1996c, 2000). The final section highlights how the new theory addresses the deficiencies of the earlier theories.

Empirical Evidence Supporting the Concept of Intuition

A fair amount of the evidence in the literature on intuition is anecdotal, and it is important to establish the experimental validity of the phenomenon before engaging in a discussion of the merits of the candidate theories. Without any doubt, the domain providing most experimental data is chess.

There is good evidence that strong players search the problem space selectively, homing in rapidly on the important moves. Klein, Wolf, Militello, and Zsombok (1995) found that the first move generated by players was usually good enough, a result that has been recently replicated in handball (Johnson & Raab, 2003). Campitelli and Gobet (2004) found that a chess grandmaster was able to correctly solve nearly 50% of problem situations within 10 s, compared to less than 5% for a weak club player. It has also been shown that performance in speed chess, where there is only about 5 s per move on average, shares 81% of the variance with the ratings based on standard chess, where players have about 180 s per move on average (Burns, 2004). The skill effect with briefly-presented chess positions (Chase & Simon, 1973; De Groot, 1965) can also be seen as a signature of intuition, in particular when one considers that masters show nearly perfect recall with a

presentation as short as 5 s and that, even though their task is to memorise the position, they also understand its meaning fairly well at the end of the presentation. In addition, eye-movement recordings during the brief presentation of the position show that masters typically look rapidly at the key elements (De Groot & Gobet, 1996).

Empirical support for the role of intuition exists in other domains as well. Thinking-aloud protocols with physics experts (Larkin, Mc Dermott, Simon, & Simon, 1980) show that they can solve routine problems in a matter of seconds. Fire-fighter commanders facing high-risk situations use intuition to make decisions under considerable time pressure (Klein, 1998). In many cases, they quickly adopt the appropriate behaviour without even considering alternatives. A similar type of behaviour has been reported by other experts, including battle commanders (Klein, 1998), managers (Patton, 2003), and intensive-care nurses (Benner, Tanner, & Chesla, 1996; Crandall & Getchell-Reiter, 1993).

Dreyfus's Theory of Expertise and Intuition

In his book *What Computers Can't Do* Dreyfus (1972) developed a wide-ranging critique of the symbolic approach in artificial intelligence, as exemplified for example by the work of Newell and Simon (1972), Minsky (1977), and McCarthy (1968). One key argument in Dreyfus's critique was that human cognition is embodied, situated, and experiential. Another key argument was that, in contradiction to classical artificial intelligence and cognitive psychology, humans do not use symbols, but perceive their environment and make decisions using holistic processes. In particular, holistic processing is characteristic of individuals that are experts in a domain. Dreyfus, a philosopher, was more interested in providing a critique of AI

based on phenomenology than in offering a detailed scientific theory, and did not develop his view in great detail, nor support it by experimental data.

In another influential book, Dreyfus (Dreyfus & Dreyfus, 1988) elaborated this view and described the steps that the aspiring expert has to go through (see also Dreyfus & Dreyfus, 1984; Dreyfus & Dreyfus, 1996, 2005). In the “novice” stage, information is acquired through instruction; domain-specific facts, features, and actions are learnt. Rules are “context-free,” in the sense that their application ignores what else is happening in the environment. The “advanced beginner” stage is attained only after substantial concrete experience with the domain. Situational elements—that is, elements that depend on the context—become meaningful and are used. In the “competence” stage, decision-making procedures are organised hierarchically. While this stage is characterized by an increased level of efficiency, planning is still to a considerable extent conscious and deliberate. In the “proficiency” stage, certain features will be perceived as salient while others will be ignored. Proficient individuals, while able to “intuitively organize and understand” the problem situations, still use analytical thinking to decide what to do next. In the final, “expertise” stage, both understanding of the task and deciding what to do is intuitive and fluid. In routine situations, “experts don’t solve problems and don’t make decisions; they do what normally works” (Dreyfus & Dreyfus, 1988, pp. 30-31). Dreyfus and Dreyfus use mostly anecdotal evidence and references to the reader’s experience to buttress their theory. In the domain of nursing, Benner and her colleagues (Benner, 1984) offer some direct empirical support in favour of the theory, based on group interviews, detailed observations, and intensive personal history interviews. (See Gobet & Chassy (in press) for a discussion of nursing expertise in the light of some of the ideas discussed in the present article.)

Dreyfus and Dreyfus have considered three ways in which the brain could produce intuitive behaviour based on experiences. In the first edition of *Mind over Machine*, they speculated that the brain could be seen as a holographic pattern recognizer. This idea was dismissed in the preface of the second edition of the book. Instead, these authors considered the possibility of using the mechanisms proposed by neural net research. This possibility was in turn dismissed in the preface of the 1992 editions of *What Computers Still Can't Do*: "It looks likely that the neglected and then revived connectionist approach is merely getting its deserved chance to fail" (Dreyfus, 1992, p. xxxviii). The final possibility considered was the approach of reinforcement learning (e.g., Tesauro, 1992), but it was concluded that this approach also met with serious practical and theoretical problems (Dreyfus, 1992), although S. E. Dreyfus (2004) provides a more optimistic evaluation.

While we have centred on Dreyfus's approach, we should mention that other authors have emphasized that intuition requires holistic processing. For example, for chess, the domain discussed at length by Dreyfus (1972) and Dreyfus and Dreyfus (1988), one can mention the proposals by De Groot (1986; see also De Groot & Gobet, 1996, for an extended discussion of De Groot's view) and Linhares (2005). At the descriptive level, the theory is in line with cognitive theories proposing that novices start with verbal, analytic knowledge and slowly move to levels where knowledge becomes unconscious (Anderson, 1982; Cleveland, 1907).

Although Dreyfus and Dreyfus's account has face value validity, it also conflicts with a fair amount of empirical data. First, there is evidence that, in many domains, expertise does not imply a decrease of abstract thought and a concomitant increase in concrete thought, as proposed by Dreyfus and Dreyfus. One of the best examples is physics, where experts in fact solve problems at a deep, abstract level,

while novices perform at a superficial, concrete level (Chi, Feltovich, & Glaser, 1981; Larkin et al., 1980).

Second, the presence of stages in expertise development is poorly documented. In addition to the well known difficulty of empirically establishing the reality of stages (van der Maas & Molenaar, 1992), there is clear evidence that individuals may be experts in one sub-field whilst performing less fluidly in another sub-field of the same domain (Benner, 1984; Gruber & Strube, 1989; Rikers et al., 2002). Although Dreyfus and Dreyfus (1988) acknowledge that the level of expertise of one individual may vary for different problems with the same area, this would suggest that the notion of stage must not be taken literally, but only suggestively. But this seems to undermine one of the main theoretical contributions of the model.

Third, a tenet of the theory, and of Dreyfus's earlier work, is that intuition is necessary for performing at expert level in what Dreyfus (1972) calls "complex formal" and "nonformal" intelligent activities and that, being analytic, heuristic-search computer programs cannot reach this level of performance (Dreyfus, 1972; Dreyfus & Dreyfus, 1988, e.g., Table 1.1; , 2005). (This is discussed in detail in the conclusion chapter of *What Computers Can't Do* (Dreyfus, 1972); see in particular the discussion surrounding Table 1 in the Conclusion chapter.) Recent developments in computer board games, for example in chess where world champion Kasparov was beaten by Deep Blue (Campbell, Hoane, & Hsu, 2002)³ and in Othello where world champion Murakami was beaten by Logistello (Buro, 1999), show that programs

³ There has been some (unsubstantiated, in our view) suggestion that Deep Blue received unfair help from its programmers. However, more recent matches pitting world champions against computer programmes running on standard PCs have consistently demonstrated that the best human players struggle against computers (see for example the 4-2 defeat of world champion Vladimir Kramnik against Deep Fritz in December 2006).

using heuristic search—without any holistic understanding of positions—can perform at very high levels (see also Strom & Darden, 1996, for a similar point). Indeed, chess grandmasters are often baffled by how their intuitions can be proven false by commercially-available computer programs.⁴

It could actually be the case that computer programs may help develop a much better understanding of chess than humans have been able to achieve (Gobet, 1993). Jansen (1992; , 1992) compared human play with endgame databases and found that even grandmasters perform weakly in simple endings. Consider the endgame King-Queen vs. King-Rook, an endgame that textbooks consider as elementary and to which they devote just a few pages. Jansen found that even world-class grandmasters made so many errors that it took, on average, four times longer than the optimal line of play to win the game. In many cases, they would have achieved only a draw instead of a win.

Fourth, while Dreyfus and Dreyfus (1988) recognize that even individuals at the expert level may need to carry out analytic problem solving, they do not supply details about how the information provided by holistic intuition may be used, for example, to guide look-ahead search in a game such as chess. In addition, the role of conscious problem solving is clearly underestimated in the theory. Based on an informal experiment with a chess international master who “more than held his own” against “a slightly weaker, but master level, player” in spite of having to add dictated numbers, Dreyfus and Dreyfus (1988, p. 33) conclude that players at the expert stage

⁴ A good example of this is the last game of the match Kramnik vs. Deep Fritz, mentioned in the previous footnote, where a series of Deep Fritz’s manoeuvres that grandmasters commenting on the game originally found primitive and naïve turned out to have deep strategic implications. Deep Fritz won the game.

can still produce “fluid and coordinated play” in spite of being “deprived of the time necessary to see problems and construct plans.” Unfortunately, not enough details are provided in Dreyfus and Dreyfus’s book to evaluate this experiment; in particular, it is unclear as to the difference in skill between the two players, whether there were behavioural differences between normal play and play with the interfering task, and, indeed, to what extent the second player was affected by the experimental setting. Well-controlled experiments with large samples (Robbins et al., 1995) have shown that a concurrent task interfering with what Baddeley (1986) calls the central executive substantially impairs the quality of the moves chosen. Robbins et al. used *tactical* chess positions, and one could argue that Dreyfus and Dreyfus’s point was that their master won his game solely through intuitive *strategic* play (roughly, position estimation and long-range planning), without using tactical play (roughly, short-term precise calculations based on thinking ahead). This seems unlikely to us, as nearly every game at master level contains moments where tactics become crucial. In addition, recent research (Chabris & Hearst, 2003; Gobet & Simon, 1996b) has clearly established that reducing thinking decreases playing skill, although one should emphasise that the level of play with grandmasters is still fairly high, and thus that some kind of pattern recognition must be involved. Thus, our disagreement is not about the importance of intuitive play at expert level, but about Dreyfus and Dreyfus’s neglect of analytical thinking—in chess, look-ahead search.

Finally, evidence from neuroscience does not support the notion of holistic pattern recognition. There is now good evidence that perception proceeds sequentially, engaging specialized modules, as is shown for example by Eimer (2000) for data on face perception and O’Rourke and Holcomb (2002) for data on word perception.

Simon's Standard Theories of Expertise and Intuition

While Dreyfus's approach is philosophical and the evidence used to support it mostly anecdotal, Simon's emphasis is on mechanistic explanations of empirical phenomena, with direct recourse to experimental data. The starting point of Simon's analysis is that experts suffer from the same cognitive limits as novices (Chase & Simon, 1973). In particular, they can pay attention to only one thing at a time, and their short-term memory (STM) is limited to just a few items. In addition, experts essentially use the same problem solving methods as novices, such as means-end analysis, progressive deepening, and use of heuristics to cut the search space down. What happens during the path from novice to expert is that individuals learn a large number of perceptual patterns that get associated with possible actions; in other words, they learn a large number of "productions" (i.e., condition-action pairs; Chase & Simon, 1973; Newell & Simon, 1972). For example, a chess player may learn that, given a certain king's side pawn structure, an attack including the sacrifice of a bishop should be considered. This chunking process is not unique to expertise, but is one basic learning mechanism found in other domains, such as verbal learning (Simon & Feigenbaum, 1964). Intuition can then be explained by the firing of a production: a pattern similar to one learned during previous experience is recognized, and thus a solution is automatically accessed (see Fig. 1).⁵ While this solution was obtained through slow problem solving mechanisms in the first instance, it is now accessed automatically by memory lookup. To some extent, within Simon's framework, intuition is just one method among others to reduce the search space.

⁵ Note that the proposed move is not necessarily the best one in a specific context—just a move that is often good in similar contexts. In the example of Figure 1, the move Ne5-g4 was actually played.

Insert Fig. 1 about here

While Simon sometimes relied on anecdotal evidence, he also used a number of experimental data to support this explanation of intuition. Data from chess (Chase & Simon, 1973) show that strong players perceive the board as chunks of pieces, and not as individual pieces, and also that they chunk sequences of moves. The chunking of actions (moves) is also apparent in learning simple puzzles such as the tower of Hanoi (Anzai & Simon, 1979). Data from physics clearly show that experts can recognize the solution of routine problems almost instantly and that, at least with routine problems, as expertise develops the search strategy changes from backward search to forward search or even forward execution—that is, proceeds through the solution with minimal search (Larkin et al., 1980).

Another source of support for Simon's theory comes from computer simulations, which establish that the mechanisms postulated by the theory are sufficient to produce the behaviour to explain. Relevant simulations, using production systems, include modelling how a novice becomes proficient in solving the tower of Hanoi (Anzai & Simon, 1979) and how backwards search is replaced by forward search as novices become experts in physics (Larkin et al., 1980). Indirect support is also offered by simulations of memory recall tasks showing how chunks—an essential component of pattern recognition and thus of intuition—are acquired in chess (Simon & Gilmarin, 1973).

In line with Simon's views, a number of theories explain intuition as recognition of perceptual patterns linked to actions, which compile domain-specific

experience acquired over years of practice and study. Among the most influential, one can mention those of Newell (1990), Saariluoma (1995), and Klein (2003).

Simon's theory of intuition has been criticized on several grounds. Dreyfus and Dreyfus (1988) note that, as chess positions are comprised of several chunks, several moves will be proposed; however, there is no provision of mechanisms explaining how only one move is selected. In addition, the types of chunk proposed by Simon are defined in isolation to other aspects of the situation. By contrast, Dreyfus and Dreyfus (1988, p. 210) argue that the position is stored as "an unanalyzable whole." They also criticize Simon's "information processing assumption that intelligence consists in drawing conclusions using features and rules," noting that high levels of expertise "are characterized by a rapid, fluid, involved kind of behavior that bears no apparent similarity to the slow, detached reasoning of the problem-solving process" (Dreyfus & Dreyfus, 1988, p. 27).⁶ De Groot (1986) argues that intuition is more than pattern recognition, emphasising its constructive and productive aspects. That is, intuition does not only reproduce previous solutions, but creatively combines elements to produce new solutions.

Holding (1985) provides additional criticisms, more aimed at Simon's general theory of expertise than at his theory of intuition in particular. Two of these criticisms are especially important theoretically: encoding into long-term memory (LTM) is faster than proposed by the chunking theory, and the size of chunks is too small to reflect conceptual knowledge and provide useful information in problem-solving situations. A third criticism—that pattern recognition is not a sufficient explanation of skill, because it applies only to the initial problem situation and does not link to

⁶ But note that this criticism appears to ignore Simon's work on chunking and pattern recognition (e.g., Chase & Simon, 1973).

look-ahead analysis—is much weaker, as Chase and Simon (1973) made it clear that pattern recognition occurs not only in the initial problem situation, but also in the problem states generated during look-ahead search.

To these criticisms, we can add that, while Simon's computer models were remarkable and insightful in their own right, they either failed to reach high levels of expertise or did so only with considerable hand-coded knowledge but no real learning. Finally, the links between intuition and emotions are not spelled out in any detail.

A New Theory of Intuition

Our discussion of Dreyfus's and Simon's theories has highlighted the features that a successful theory of expert intuition should have: it should explain the rapid onset of intuition and its links with emotion, provide mechanisms for learning, have processes showing how perception is linked to action, and explain how experts capture the entirety of a situation. In this section, we develop such a theory, taking as basis the template theory of memory developed by Gobet and Simon (Gobet & Simon, 1996c, 2000).

The template theory was developed to correct a number of weaknesses of the chunking theory (Chase & Simon, 1973), of which it is a modification and extension. These include the fact that players use larger chunks than those proposed by the chunking theory, the failure of Simon and Gilmarin's (1973) computer simulations of memory recall to reach master level and the fact that the chunks learnt were pre-selected by the programmers, and weaknesses in the way PERCEIVER (Simon & Barenfeld, 1969) accounted for chess players' eye movements. Aspects of the template theory are implemented in a computer program known as CHREST (Chunk Hierarchy and REtrieval STructures) (Gobet & Simon, 2000; Gobet & Waters, 2003), and the fact that CHREST simulates not only the phenomena tackled by Simon and

Gilmartin's (1973) and Simon and Barenfeld (1973), but also a substantial number of new phenomena (see section "Simulations with the CHREST model" below) clearly shows that the weaknesses of the earlier programs have indeed been corrected in the new theory without inadvertently creating new problems.

Overview of the Theory

We carefully distinguish between the features of the template theory that have been implemented in CHREST and in other programs,⁷ those that are part of the theory but have not been implemented yet, and those that we have added to the theory to account for the link between intuition and emotions.

The CHREST model.

Components. Like the original chunking theory, template theory proposes that expertise is made possible by the acquisition of a large number of chunks, some of which are linked to possible actions. A key addition of template theory is the assumption that some patterns that recur often in the environment give rise to chunks that develop into more complex data structures called templates. Templates are similar to schemata (Bartlett, 1932; Minsky, 1975) in that they possess both a core, made of stable information, and slots, made of variable information. Unlike previous schema theories, template theory proposes detailed mechanisms as to how templates—both their core and their slots—are acquired (see below).

In CHREST, chunks and templates (which are a special case of chunks) are indexed by a discrimination network (Simon & Gilmartin, 1973), which consists of a network of sequential tests enabling the access to information in LTM.⁸ While

⁷ The code of these programs (in Lisp) is available from the first author.

⁸ Neurally, it has been proposed that chunks and templates are implemented as cell assemblies (Chassy & Gobet, 2005).

learning is assumed to be slow (e.g., 10 s to create a new chunk), access of chunks by sorting the discrimination network is assumed to lead to fast recognition of objects (a few hundred milliseconds).

Chunks may be connected by similarity links if they have enough elements in common. In addition to the discrimination net, the model has three components: an LTM, a visual STM, and a “mind’s eye.” LTM contains chunks, productions and schemata. Visual STM has a capacity of three chunks. It is a queue, meaning that, when a new chunk enters STM where it is already full, the oldest chunk “pops out” from STM. The exception is that the largest chunk is kept in STM until a larger chunk is met. Templates, which have slots in which variable information can be stored, are a special type of chunk. Finally, the mind’s eye stores visuo-spatial information for a short time; it is the place where, for example, the trajectories of pieces are computed. The main mechanisms used by CHREST deal with eye fixations, STM management, LTM learning, and information update in the mind’s eye. In general, it is assumed that humans are conscious of the information held in STM and in the mind’s eye, but not of the information and processes used during learning and recognition.

While CHREST has been applied to other domains (Gobet et al., 2001), we focus on chess in the explanations that follow, not only because this domain provides some the best evidence for intuition (see above), but also because both Dreyfus and Simon heavily refer to chess in their theory of intuition.

Insert Fig. 2 about here

Fig. 2 illustrates the main components of the theory, with chess as the task environment. A simulated eye scans the board, and the information within the visual field is input to the discrimination network, which leads to the access of a certain node in LTM. A pointer to this node is placed in short-term memory, and the information is also unpacked in pictorial short-term memory (the “mind’s eye”).⁹ This sequence of operations is assumed to be repeated when players look at a position.

Eye fixations. CHREST’s attention is directed by eye movements. The program attempts to use information provided by the largest chunk met at any given point to fixate a location. In chess, this operation is performed by following a branch that is stored below the chunk and fixating the square associated with this branch. As an example, let us assume that this is the chunk depicted in grey in Fig. 3. CHREST would take the link leading to the most recently created node (in our example, “white pawn on f2”), and fixate on the square indicated by this link, in this case the square f2. If a white pawn is indeed located on f2, then a larger chunk has been found, and thus more information retrieved from LTM. Although this guess, informed by experience, may sometimes be incorrect, it tends to produce eye movements that are similar to those of experts. For example, it is this mechanism that leads the program to fixate on semantically-important squares in a proportion similar to experts’ (see De Groot & Gobet, 1996, for details).

If it is not possible to use this eye-movement mechanism based on knowledge (for example, because there is no branch in the discrimination net below the largest

⁹ Neurally, such pointers might be implemented by short-term memory neurons in the prefrontal cortex firing in synchrony with neurons in posterior areas of the brain; the limited capacity of STM—that is, the limited number of pointers that can be held in STM—is then a function of the number of distinct frequencies available (Ruchkin, Grafman, Cameron, & Berndt, 2003).

chunk), the program draws on alternative mechanisms such as fixation on a perceptually salient object or on a region of the display that has not been visited yet (see De Groot & Gobet, 1996). Whereas novices' eye movements are mainly directed by such heuristics, most experts' eye movements use the first mechanism and are directed by the structure of the discrimination network. An interesting feature of the theory is thus that it includes mechanisms detailing how perception determines what will be learned, on the one hand, and how learned knowledge determines what will be perceived, on the other.

Learning chunks and creating templates. After each new fixation, the model filters the information in the visual field through the discrimination net. In the chess simulations, the field of vision is limited to two squares away from the fixation point in each direction, so that a maximum of 25 squares can be perceived at any time (see De Groot & Gobet, 1996, for empirical data supporting this choice). An external pattern is encoded as a list of the pieces on their squares; for example, in Fig. 2, the pattern on the right of the position would be encoded as: (♖ f2, ♖ g2, ♖ h2, ♔g1, ♕e2, ♜f4).

Two learning mechanisms are used, familiarisation and discrimination. When a new (external) object is perceived, it is sorted through the discrimination net. When a node is reached, the object is compared with the information stored with this node, which is known as the "image." If the image under-represents the object, new features are added to the image (familiarisation). If there is a mismatch between the information in the image and the object, a new node is created below the current node by recursively adding to it some of the mismatching information (discrimination). CHREST also creates "similarity links" between nodes and templates. Each chunk arriving into STM is compared with the largest chunk already stored there. When the

two chunks are sufficiently similar, a similarity link is created between them. During the recognition phase, a similarity link can be used to move from the node reached by sorting to another similar node.

Templates are chunks that possess at least one slot where variable information can be stored. Template slots are created when enough nodes share related information below a node that is sufficiently large.¹⁰ Fig. 3 illustrates this mechanism for the domain of chess. There are four nodes below the node depicted in grey; the information “white pawn” occurs three times, and the information “square e4” occurs twice. In this simplified example, we assume that the minimum number of occurrences is two, and thus slots are created.

 Insert Fig. 3 about here

Time parameters. Each process has a time cost, which enables precise and quantitative simulations to be carried out. For example, the discrimination process, whereby a new node (and a branch leading to it) is added in the discrimination net, takes 8 s, and the familiarisation process, whereby information is added to an extant chunk, takes 2 s. Filling a template slot is faster and takes 250 msec. A full discussion of the time parameters in CHREST is provided by De Groot and Gobet (1996) and Gobet and Simon (2000).

Learning phase. Using the mechanisms just described, CHREST learns chunks and templates by scanning a large database of positions taken from master

¹⁰ In the chess simulations, the requirement is that the target node contains at least five elements and that at least three nodes below that node share identical information (either a square, a type of piece, or a chunk) (Gobet & Simon, 2000).

games, moving its simulated eye around the board, and sorting the pieces within its visual field through the discrimination network. Thus, learning is implicit, incremental and unsupervised, and it essentially captures the regularities of the environment without producing a statistical representation.

Simulation of memory experiments. During the presentation of a position, CHREST fixates on squares using the eye movement mechanisms described above. Each fixation defines a visual field (see above), and the pieces belonging to this visual field are sorted through the discrimination net. If a chunk (a pattern already familiar to the discrimination net) is found, a pointer to it is placed in STM, or, when possible, the chunk is used to fill one slot of a template. If the presentation time is long enough, the program learns using the mechanisms described above.

During the reconstruction of a position, CHREST first draws on the information stored in STM, and then information stored in LTM. Pieces are placed sequentially. If a piece has already been replaced on the board from a previous chunk, it is ignored. Conflicts can occasionally occur: for example, a square containing several pieces. Such conflicts are resolved sequentially, making use of the frequency with which each placement is suggested. It is therefore possible for the program to “change its mind” about the location of a piece or the contents of a square (see example in the appendix), and so do human players.

Extensions of CHREST for problem solving.

The idea that the recognition of patterns of chess piece allows for the accessing of information about good moves is embodied in CHUMP (CHUnks and Moves Patterns; Gobet & Jansen, 1994), a variant of CHREST. CHUMP stores two types of knowledge in two different but linked discrimination nets. The first relates to patterns of pieces (the type of chunks learned by CHREST). The second relates to

moves and sequences of moves. During learning, where positions from master games are presented, patterns of pieces are associated with moves. During the performance phase, patterns of pieces act as conditions, and moves as actions. When the recognized piece patterns suggest different moves, the program resolves the conflict by using a function that combines the number of chunks voting for a given move and the number of times the move has been seen with a given pattern during learning. The program could play chess by pure pattern recognition, but its lack of look-ahead abilities meant that its level of play was low. Another limitation of CHUMP is that it learns only a small part of the knowledge that chess experts presumably encode as productions. The literature on chess skill (e.g., Gobet, de Voogt, & Retschitzki, 2004) suggests that chess experts have other productions where the conditions consist of nodes containing information such as positional concepts, tactical features, etc., while the nodes denoting actions encode information such as plans, heuristics, tactical tricks, etc.

A stochastic model, SEARCH (Gobet, 1997) puts together several mechanisms that are proposed by the template theory but not implemented in CHREST. Unlike CHREST, the model does not carry out the detail of the postulated processes, but computes key measures, such as depth of search or the number of moves searched per minute, as a function of the number of chunks and templates. SEARCH explicitly combines pattern recognition, search, and mental imagery. It also includes assumptions about the time needed for cognitive operations, as well as assumptions about the “fuzziness” of the images kept in the mind’s eye. Chunks and templates favour deeper search, because they suggest potential moves automatically (templates also facilitate long-term memory encoding, maintenance of information in the mind’s eye, and more abstract search). On the other hand, these memory

structures favour shorter search, as they provide powerful evaluations that cut down the need for search. The net product, as shown in computer simulations, is that average depth of search follows a power function of skill—a prediction consistent with the data.

Adding emotions to the template theory.

Starting with De Groot (1965), researchers have often emphasized the role of emotion for intuition (Bechara, Damasio, Tranel, & Damasio, 1997; Benner, 1984), but this is a feature that was not covered in any detail by the chunking theory, Dreyfus's theory, nor the original version of the template theory. One important contribution of the current article is to show how the template theory can be extended to include emotions and to speculate on the biological basis of this link. The available evidence supports the view that a link associates simple cognitions and simple emotional responses (LeDoux, 1999). Similarly, there is evidence that complex representations, stored in the inferior temporal cortex, are associated to neural nets coding for reward (Rolls, 2003), and that such neural networks underlie the automatic retrieval of emotional responses (Panksepp, 1998). Cognitions, whether simple or complex, are thus associated with emotional responses. We propose that, during the activities taking place in the practice and study of a domain, chunks and templates become associated to emotional responses. Later, when a chunk or a template is retrieved from LTM, it may activate one or several emotional responses. These responses are analysed by an emotional processor that determines what emotional response is to be given priority. The emotional processor not only triggers the body changes but also instigates modulation of cognitive processing. It is worth noting that emotional responses, and thus cognitive modulation, are submitted to huge personal

variability, known as affective style (Davidson & Irwin, 1999), which may be partly explained by different histories of learning crystallised in LTM structures.

Similar to what has been shown with emotional conditioning, we propose that chunks are associated to emotional responses by hebbian learning. For example, in an adversarial game like chess, we can expect typical defensive or attacking chunks to be associated with reward or rejection. The purpose of emotional responses would be to draw the player's attention towards possible dangers in the position. This emotional bias would contribute to the selection of an appropriate option: a kind of emotionally-driven decision-orienting heuristic. As expertise develops, the alerting system made up of emotions tunes the emotional response for the chunks stored in LTM. In a later phase, the emotional system may code the emotional responses in reply to frequently-encountered combinations of chunks.

Simulations with the CHREST model

Simulations with CHREST show that the theory accounts for a wide range of data, both quantitatively and qualitatively, on skilled and unskilled chess perception, mental imagery, learning and memory, including: eye movements during the 5-s presentation of a position; memory for game positions as well as positions randomised or modified in various ways; effect of presentation time (from 1 s to 60 s); and how novices acquire chunks and templates (De Groot & Gobet, 1996; Gobet, 1993; Gobet & Jackson, 2002; Gobet & Simon, 2000; Gobet & Waters, 2003; Waters & Gobet, in press). When considering the simulations, the coverage is broader than that offered by the chunking theory, and, unlike the earlier computer model by Simon and Gilmarin (1973), CHREST actually carries out the selection of the chunks to learn, without the need of human supervision. This is achieved by simulating eye movements even during the learning phase.

To illustrate further the way CHREST works, we describe simulations for three aspects of the theory that will play an important role in explaining intuition: template formation, eye movements, and dynamic character of the simulations.

Template formation.

Given the importance of templates in the theory, it is important to show that the mechanisms used lead to the formation of templates that are plausible. Fig. 4 shows a sample of templates that were created by CHREST. For each position, the core of the template is shown by the pieces placed on the board; the slots for squares are indicated by a dot on the square, and the slots for pieces are indicated by the icons shown below each position. Based on the joint judgement of the authors, who are both chess masters, these templates correspond reasonably well to typical chess positions. Note that in these examples, templates are not only constructed for typical openings, but also for typical offensive or defensive set-ups in middle-game positions (e.g., first position in Fig. 4).

Insert Fig. 4 about here

Eye movements.

As we have seen in the description of eye fixations, one of CHREST's attractive features is that its domain-specific knowledge is used to direct a large proportion of eye movements in the simulation of experts, but without using rules. This makes possible to simulate several key aspects of masters' eye movements in chess. De Groot and Gobet (1996) have shown that the program captures the main features of human behaviour: average duration of fixations, low variability in the

duration of fixations, proportion of the board covered, and proportion of the semantically-important squares covered. For example, the average fixation duration is 272 ms ($SD = 97$ ms) for CHREST, which is in close agreement with the average duration of the human masters (260 ms; $SD = 100$ ms).

Fig. 5a shows a typical master pattern of eye movements, for the position shown in Fig. 2, and Fig. 5b shows, for the same position, a typical run of CHREST. While the exact sequences of eye fixations differ—the sequences of eye movements also differ across human players of the same skill level—it is clear that the program reproduces the key features of the human pattern.

 Insert Fig. 5 about here

Time course of constructing the internal representation of a position.

A key assumption of CHREST is that there is a close interaction between STM and LTM. This interaction is mainly made possible by three mechanisms. First, when an LTM chunk is recognized, a pointer to this chunk is stored into STM. Second, the information in STM may be used for further learning. Third, when a template is held in STM, information can rapidly be added to its slots. Appendix 1 illustrates some of these ideas with a detailed run of CHREST in a recall task, where the position is presented for five seconds; again, the position depicted in Fig. 2 is used.

Intuition

We are now in a position to discuss the contribution of the template theory to our understanding of intuition. In this respect, the template theory shares several features with the chunking theory, including the assumption that intuition can be

largely explained by pattern recognition; that chunks, which are learnt implicitly, mediate pattern recognition; that chunks give access to information about what kinds of action can be executed; and that there is close interplay between pattern recognition and search, with the implication that intuition affects the entire decision process, not only its early phase. In this respect, both theories agree with De Groot's statement (1986, p. 70) that "... intuitive processing is omnipresent in human thinking."

There are also important differences between the old and new theory. As we shall see in the next section, these novel features are crucial for explaining key aspects of intuition. These include the presence of similarity links between nodes in the discrimination net; more complex data structures (templates) in LTM; and provision of mechanisms for incrementally creating templates and automatically linking actions to perceptual patterns. In particular, the presence of templates enables internal representations of the environment to be constructed at a higher level of abstraction than assumed in the chunking theory, while still explaining the speed at which these representations are created. Another important novel feature of the theory is that it closely links attention, perception, learning, and action, in that it proposes mechanisms showing how LTM knowledge—in this case, the structure of the discrimination net—directs eye movements; this provides a powerful explanation of why the key features in a scene are generally perceived rapidly by experts. Finally, the extended theory accounts for how emotions affect cognition during learning and performance.

How the New Theory Addresses the Issues Problematic for Dreyfus's and Simon's Theories

Our discussion of Dreyfus's and Simon's theories has led to the identification of problems in both of them, problems that relate to deep issues in our understanding

of intuition. If the extended template theory is a valid theory of intuition, it should be able to address these issues satisfactorily. It is therefore important to review these questions from the point of view of template theory.

Holistic vs. Local Processing

Empirical data suggest that experts process information at various levels of granularity, including low-level features and high-level representations. In particular, the recall of game and random chess positions has shed important light on this issue, and has direct relevance to the question as to whether intuition is always holistic in nature. CHREST accounts for recall data obtained with brief presentation times, simulating data such as the percentage of correctly recalled pieces, the type and number of errors, as well as the size and number of chunks. For example, the program replicates how players of different skill levels perform with presentation times ranging from 1 to 60 s, and in particular the rapid improvement shown by masters with game positions after 1 s (Gobet & Simon, 2000). As noted by Gobet and Simon, this phenomenon is directly affected by the presence of templates, and in particular the assumption that encoding in the template slots is rapid once a template has been recognized (see also the discussion of the CHREST trace presented in the appendix).

Another important result—directly addressing the issue of local processing—is that CHREST also accounts for the small skill effect present in the recall of *random* positions. While this effect is not as large as with game positions, it is reliable and has been replicated several times (Gobet & Simon, 1996a, 2000). The program accounts for this effect by recognizing local patterns that show up serendipitously even in random positions. The larger the discrimination net, the higher the probability of finding chunks for such patterns, hence the skill effect. It is unclear how a holistic

theory such as Dreyfus's can account for these data, as they seem to rely on processing local aspects of the positions. In general, the pattern of eye movements during the 5-s presentation of a position also supports the hypothesis of a progressive and serial construction of an internal representation rather than holistic processing (De Groot & Gobet, 1996; Simon & Barenfeld, 1969). Thus, the evidence seems to point to large, "holistic" representations being constructed by local mechanisms.

From Abstract to Concrete?

An important prediction of Dreyfus and Dreyfus's theory (1988) is that, as novices become experts, there is a transition from analytic to intuitive, from abstract to concrete knowledge. We have seen earlier that, at least in some domains of expertise and in some tasks, the opposite pattern is actually observed. But this also seems to be a simplified picture. As shown by research in domains such as chess (De Groot, 1965), physics (Larkin et al., 1980; Simon & Simon, 1978), and nursing (Benner, 1984; Gobet & Chassy, in press), the pattern of learning is more complex and incorporates a progression from analytic to intuitive knowledge but also an increased ability to deal with abstractions. Thus, an expert in physics will both recognize concrete patterns rapidly and understand the problems at a higher level of abstraction than a novice. The template theory readily deals with the acquisition of different types of knowledge and representation (e.g., diagrammatic and algebraic), as has been discussed at length in the context of education (Gobet, 2005; Gobet & Wood, 1999). The theory predicts that perceptual, schematic, and procedural as well as concrete and abstract knowledge are acquired in parallel, and thus that these types of knowledge should overlap with experts.

Analytic and Intuitive Behaviour

Just like Simon's theory, our theory has the advantage over Dreyfus's that the links between intuition and slower problem solving behaviour are made explicit. The key idea, already present in Simon's earlier work and fully developed in Gobet (1997), is that problem solving involves cycles interleaving pattern recognition and search. At the beginning, there is an attempt to access a chunk or a template in LTM. The more expert the individual, the more likely this will be successful. If a chunk or a template is accessed, the information linked to it is used to carry out further searching of the problem, and this cycle continues. In cases where no chunk or template can be found, or where no information is associated to them, weaker heuristics are used, either domain-specific heuristics or domain-general heuristics such as means-end analysis. When the problem is easy, the correct solution can be retrieved by LTM look-up.

As noted above, one of the objections filed by Dreyfus and Dreyfus (1988) to Simon's idea of intuition as pattern recognition was that there was no explanation as to how a single action (a move, in the case of chess) could be chosen while several chunks could be identified. The situation is actually worse, as a single chunk could propose several actions. This is a standard issue with production systems, technically known as conflict resolution, and several solutions have been proposed (Neches, Langley, & Klahr, 1987; see also the description of CHUMP above).

Due to limits in existing technology in the sixties and seventies, Simon could not develop simulation programs able to show that such pattern-recognition mechanisms could indeed lead to the selection of a move and that the pattern-action pairs could be learned automatically. As we have seen earlier, the CHUMP program (Gobet & Jansen, 1994) is doing just this for the domain of chess. With respect to

intuition, it is of particular interest that this program performed better in situations requiring a “positional judgment” than in tactical positions, where look-ahead search becomes more critical. Positional judgment in chess is often presented as a paradigmatic example of intuition, not only by Dreyfus and Dreyfus (1988), but also in the popular literature on chess (Kotov, 1971).

The Problem of Small Chunks

As noted above, two other objections to Simon’s theory are that the chunks identified by Chase and Simon (1973) for chess may be too small to elicit moves and that they do not capture the whole of the position. The simulations carried out by Gobet and Jansen (1994) show that the first objection does not apply. As to the second objection, there is indeed ample evidence that strong players use high-level representations at a more abstract level than the piece locations encoded by chunks and that, at least in some cases, they perceive the entire board as a single unit. This evidence includes the analysis of verbal protocols in problem-solving tasks (De Groot, 1965), recall tasks (De Groot, 1965; De Groot & Gobet, 1996), and classification tasks (Freyhoff, Gruber, & Ziegler, 1992). As noted above, the template theory captures this aspect of expert perception in chess. As a matter of fact, postulating structures that could potentially cover the entire problem situation and that have schema-like properties was one of the motivations behind the development of the template theory. Although templates capture the “wholeness” of perception taken by Dreyfus to be a signature of expertise, their construction is incremental, with larger chunks being recursively produced by the conjunction of smaller chunks.

Conclusion

In this paper, we have briefly considered the empirical data supporting the concept of intuition, before discussing two influential theories of intuition, that of

Hubert Dreyfus and that of Herbert Simon. We have noted that, ironically, Dreyfus uses experts' intuition as one of the main grounds for which information processing psychology (and classical artificial intelligence) is doomed to fail, ignoring empirical and theoretical work by Simon and others showing that simple information processing mechanisms might explain this phenomenon. We have also noted that chess has often been used to illustrate the putative bankruptcy of rule-based and symbolic thinking. By contrast, the empirical evidence we have discussed has illustrated situations where these symbolic techniques do better than human intuition.

Our critical analysis of the two theories has established that, while both address important aspects of expert intuition, both fail to account for the empirical data thoroughly. To address this theoretical gap, we have shown how the template theory, a modification of the chunking theory, accounts for most of the empirical data linked to intuition. In addition to pattern recognition—already present in the chunking theory—the key mechanisms relate to the interaction between perception, attention, and learning, and to the creation and use of templates. These schema-like structures enable information to be encoded both rapidly and at a high level of representation. A further important addition consisted of mechanisms linking chunks and templates to emotions.

This paper has emphasised the differences between the three theories, but it is fair to acknowledge that they share a number of similarities: beyond accepting intuition as a genuine phenomenon, all three theories emphasise the essential role of perception, the fluid, automatised, and rapid behaviour characteristic of experts' intuition, and the long time required to become an expert. They also all stress the importance of discrimination and association in explaining experts' behaviour,

although the holistic nature of these processes, essential in Dreyfus's theory, is not shared by the other two theories.

The differences between the three theories have implications for practice and research. The assumption that experts' knowledge is composed of chunks, as opposed to Dreyfus' assumption that it is holistic in nature, makes it possible to design curricula where the instructional material is decomposed into small bits and where computer-based tutors may be used (Anderson, Reder, & Simon, 2000; Gobet & Wood, 1999). Assuming that knowledge also consists of templates leads to considerations as to how schematic knowledge can be best acquired and taught (Gobet, 2005); for example, variety in the curriculum material is an essential requirement for making possible the acquisition of templates. By contrast, an emphasis on the holistic nature of expertise, with the implication that experts' understanding cannot be analysed into components, leads to different types of curricula, where engagement in real-life situations is emphasized. The importance of such situation is of course not negated by chunk-based approaches, but seen as complementary to other instructional methods. As for empirical research, the impact of the two approaches can readily be seen. Traditional research on expertise has been largely motivated by Chase and Simon's (1973) chunking theory, and has been characterised by a substantial number of experimental and quantitative observational studies, and to a lesser extent computer models (for reviews, see Ericsson, Charness, Feltovich, & Hoffman, 2006; Gobet, de Voogt, & Retschitzki, 2004). Research on nursing expertise, perhaps the domain that has been most influenced by Dreyfus' theory, is mostly made up of qualitative observational studies, of which many consist of phenomenological analysis (e.g., Benner, 1984; Benner et al., 1996), and quantitative approaches are explicitly considered suspicious (Benner, 1984).

In sum, this paper has presented a new theory of intuition, part of which is formally expressed as a computer program. We have argued that it accounts for all the phenomena taken as signatures of intuition. Crucially, the new theory leads to the conclusion that, while aspects of expert intuition can be characterized as holistic, the mechanisms that lead to them are local.

References

- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89, 396-406.
- Anderson, J. R., Reder, L. M., & Simon, H. A. (2000, Summer). Applications and misapplications of cognitive psychology to mathematics education. *Texas Education Review*.
- Anzai, Y., & Simon, H. A. (1979). The theory of learning by doing. *Psychological Review*, 86, 124-140.
- Baddeley, A. (1986). *Working memory*. Oxford: Clarendon Press.
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. New-York: Cambridge University Press.
- Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, 275, 1293-1295.
- Benner, P. (1984). *From novice to expert: Excellence and power in clinical nursing practice*. Menlo Park, CA: Addison-Wesley.
- Benner, P., Tanner, C., & Chesla, C. (1996). *Expertise in nursing practice: Caring, clinical judgment, and ethics*. New York: Springer Publishing.
- Burns, B. D. (2004). The effects of speed on skilled chess performance. *Psychological Science*, 15, 442-447.
- Buro, M. (1999). How machines have learned to play Othello. *IEEE Intelligent Systems*, 14, 12-14.
- Campbell, M., Hoane, A. J., & Hsu, F. H. (2002). Deep Blue. *Artificial Intelligence*, 134, 57-83.
- 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.
- Chabris, C. F., & 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.

- Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215-281). New York: Academic Press.
- Chassy, P., & Gobet, F. (2005). A model of emotional influence on memory processing. In L. Cañamero (Ed.), *AISB 2005: Symposium on agents that want and like*. Hatfield, UK: University of Hertfordshire.
- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121-152.
- Cleveland, A. A. (1907). The psychology of chess and of learning to play it. *The American Journal of Psychology*, XVIII, 269-308.
- Crandall, B., & Getchell-Reiter, K. (1993). Critical decision method: A technique for eliciting concrete assessment indicators from the "intuition" of NICU nurses. *Advances in Nursing Science*, 16, 42-51.
- Davidson, R. J., & Irwin, W. (1999). The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Sciences*, 3, 11-21.
- De Groot, A. D. (1965). *Thought and choice in chess (first Dutch edition in 1946)*. The Hague: Mouton Publishers.
- De Groot, A. D. (1986). Intuition in chess. *Journal of the International Computer Chess Association*, 9, 67-75.
- De Groot, A. D., & Gobet, F. (1996). *Perception and memory in chess: Heuristics of the professional eye*. Assen: Van Gorcum.
- Dreyfus, H. L. (1972). *What computers can't do: A critique of artificial reason*. New York, NY: Harper & Row.
- Dreyfus, H. L. (1992). *What computers still can't do: A critique of artificial reason*. Cambridge, MA: The MIT Press.
- Dreyfus, H. L., & Dreyfus, S. E. (1984). From Socrates to expert systems: The limits of calculative rationality. *Technology in Society*, 6, 217-233.
- Dreyfus, H. L., & Dreyfus, S. E. (1988). *Mind over machine: The power of human intuition and expertise in the era of the computer* (2nd ed.). New York: Free Press.
- Dreyfus, H. L., & Dreyfus, S. E. (1996). The relationship of theory and practice in the acquisition of skill. In P. Benner, C. Tanner & C. Chesla (Eds.), *Expertise in nursing practice: Caring, clinical judgment, and ethics* (pp. 29-47). New York: Springer Publishing.

- Dreyfus, H. L., & Dreyfus, S. E. (2005). Expertise in real world contexts. *Organization Studies*, 26, 779-792.
- Dreyfus, S. E. (2004). Totally model-free learned skillful coping. *Bulletin of Science, Technology & Society*, 24, 182-187.
- Eimer, M. (2000). Event-related brain potentials distinguish processing stages involved in face perception and recognition. *Clinical Neurophysiology*, 111, 694-705.
- Ericsson, K. A., Charness, N., Feltovich, P. J., & Hoffman, R. R. (2006). *The Cambridge handbook of expertise and expert performance*. New York, NY: Cambridge University Press.
- Feigenbaum, E. A., & Simon, H. A. (1984). EPAM-like models of recognition and learning. *Cognitive Science*, 8, 305-336.
- Freyhoff, H., Gruber, H., & Ziegler, A. (1992). Expertise and hierarchical knowledge representation in chess. *Psychological Research*, 54, 32-37.
- Gobet, F. (1993). *Les mémoires d'un joueur d'échecs [Chess players' memories]*. Fribourg: Editions universitaires.
- Gobet, F. (1997). A pattern-recognition theory of search in expert problem solving. *Thinking and Reasoning*, 3, 291-313.
- Gobet, F. (2005). Chunking models of expertise: Implications for education. *Applied Cognitive Psychology*, 19, 183-204.
- Gobet, F., & Chassy, P. (in press). Two theories of expert intuition in nursing: A discussion paper. *International Journal of Nursing Studies*.
- Gobet, F., de Voogt, A. J., & Retschitzki, J. (2004). *Moves in mind: The psychology of board games*. Hove, UK: Psychology Press.
- Gobet, F., & Jackson, S. (2002). In search of templates. *Cognitive Systems Research*, 3, 35-44.
- Gobet, F., & Jansen, P. J. (1994). Towards a chess program based on a model of human memory. In H. J. van den Herik, I. S. Herschberg & J. W. H. M. Uiterwijk (Eds.), *Advances in Computer Chess 7* (pp. 35-60). Maastricht: University of Limburg Press.
- Gobet, F., & Simon, H. A. (1996a). Recall of rapidly presented random chess positions is a function of skill. *Psychonomic Bulletin & Review*, 3, 159-163.

- Gobet, F., & Simon, H. A. (1996b). The roles of recognition processes and look-ahead search in time-constrained expert problem solving: Evidence from grandmaster level chess. *Psychological Science*, *7*, 52-55.
- Gobet, F., & Simon, H. A. (1996c). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology*, *31*, 1-40.
- Gobet, F., & Simon, H. A. (2000). Five seconds or sixty? Presentation time in expert memory. *Cognitive Science*, *24*, 651-682.
- Gobet, F., & Waters, A. J. (2003). The role of constraints in expert memory. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *29*, 1082-1094.
- Gobet, F., & Wood, D. J. (1999). Expertise models of learning and computer-based tutoring. *Computers and Education*, *33*, 189-207.
- Gruber, H., & Strube, G. (1989). Zweierlei Experten : Problemisten, Partierspieler und Novizen bei Lösen von Schachproblemen. *Sprache & Kognition*, *8*, 72-85.
- Holding, D. H. (1985). *The psychology of chess skill*. Hillsdale, NJ: Erlbaum.
- Jansen, P. J. (1992). KQKR - Awareness of a fallible opponent. *ICCA Journal*, *15*, 111-131.
- Jansen, P. J. (1992). *Using knowledge about the opponent in game-tree search*. Unpublished Ph. D. CMU-CS-92-192, Carnegie Mellon, Pittsburgh.
- Johnson, J. G., & Raab, M. (2003). Take The First: Option-generation and resulting choices. *Organizational Behavior and Human Decision Processes*, *91*, 215-229.
- Jongman, R. W. (1968). *Het oog van de meester*. Assen: Van Gorcum.
- Klein, G. A. (1998). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.
- Klein, G. A. (2003). *Intuition at work*. New York, NY: Currency and Doubleday.
- Klein, G. A., Wolf, S., Militello, L., & 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.
- Larkin, J. H., Mc Dermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, *208*, 1335-1342.
- LeDoux, J. (1999). *The emotional brain*. London, UK: Phoenix.
- Linhares, A. (2005). An active symbols theory of chess intuition. *Minds and Machines*, *15*, 131-181.

- McCarthy, J. (1968). Programs with common sense. In M. Minsky (Ed.), *Semantic information processing* (pp. 403-418). Cambridge, MA: MIT Press.
- Minsky, M. (1975). A framework for representing knowledge. In P. H. Winston (Ed.), *The Psychology of computer vision* (pp. 211-277). New-York: McGraw-Hill.
- Minsky, M. (1977). Frame-system theory. In P. N. Johnson-Laird & P. C. Wason (Eds.), *Thinking. Readings in Cognitive Science* (pp. 355-376). Cambridge: Cambridge University Press.
- Neches, R., Langley, P., & Klahr, D. (1987). Learning, development, and production systems. In D. Klahr, P. Langley & R. Neches (Eds.), *Production system models of learning and development* (pp. 1-53). Cambridge, MA: MIT Press.
- Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- O'Rourke, T. B., & Holcomb, P. J. (2002). Electrophysiological evidence for the efficiency of spoken word processing. *Biological Psychology*, *60*, 121-150.
- Panksepp, J. (1998). *Affective neuroscience*. Oxford, UK: Oxford University Press.
- Patton, J. R. (2003). Intuition in decisions. *Management Decision*, *41*, 989-996.
- Reynolds, R. I. (1982). Search heuristics of chess players of different calibers. *American Journal of Psychology*, *95*, 383-392.
- Rikers, R. M. J. P., Schmidt, H. G., Boshuizen, H. P. A., Linssen, G. C. M., Wesseling, G., & Paas, F. G. W. C. (2002). The robustness of medical expertise: Clinical case processing by medical experts and subexperts. *American Journal of Psychology*, *115*, 609-629.
- Robbins, T. W., Anderson, E., Barker, D. R., Bradley, A. C., Fearnlyhough, C., Henson, R., et al. (1995). Working memory in chess. *Memory and Cognition*, *24*, 83-93.
- Rolls, E. T. (2003). Vision, emotion and memory: From neurophysiology to computation. *International Congress Series*, *1250*, 547-573.
- Ruchkin, D. S., Grafman, J., Cameron, K., & Berndt, R. S. (2003). Working memory retention systems: A state of activated long-term memory. *Behavioral and Brain Sciences*, *26*, 709-+.
- Saariluoma, P. (1995). *Chess players' thinking: A cognitive psychological approach*. London: Routledge.

- Simon, D. P., & Simon, H. A. (1978). Individual differences in solving physics problems. In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp. 323-348). Hillsdale, N.J: Erlbaum.
- Simon, H. A. (1989). *Models of thought (vol II)*. New Haven, CT: Yale University Press.
- Simon, H. A., & Barenfeld, M. (1969). Information processing analysis of perceptual processes in problem solving. *Psychological Review*, 7, 473-483.
- Simon, H. A., & Feigenbaum, E. A. (1964). An information processing theory of some effects of similarity, familiarity, and meaningfulness in verbal learning. *Journal of Verbal Learning and Verbal Behavior*, 3, 385-396.
- Simon, H. A., & Gilmartin, K. J. (1973). A simulation of memory for chess positions. *Cognitive Psychology*, 5, 29-46.
- Strom, J. D., & Darden, L. (1996). Is artificial intelligence a degenerating program? [Review of the book *What computers still can't do*]. *Artificial Intelligence*, 80, 151-170.
- Tesauro, G. (1992). Practical issues in temporal difference learning. *Machine Learning*, 8, 257-277.
- Tikhomirov, O. K., & Poznyanskaya, E. D. (1966). An investigation of visual search as a means of analyzing heuristics. *Soviet Psychology*, 5, 2-15.
- van der Maas, H. L. J., & Molenaar, P. C. M. (1992). Stagemwise cognitive development: An application of catastrophe theory. *Psychological Review*, 99, 395-417.
- Waters, A. J., & Gobet, F. (in press). Mental imagery and chunks: Empirical and computational findings. *Memory & Cognition*.

Appendix

Trace of CHREST in a memory recall task

Fig. A1 illustrates the information held in CHREST's STM during the five-second presentation of a position (see Fig. 2) during a task, using the timing of the eye fixations as a clock. Given the context of this paper, we are mostly interested in what happens at the beginning, and thus provide all the STM states during the first two seconds. For each panel in the Fig., the first line shows the time at which the fixation was carried out, the following lines the state of STM (where #C means "chunk" and #T means "template"), and the diagram shows the pieces that would be replaced on the board if CHREST had to recall the position at this point. The pieces or squares in grey indicate that this information has been encoded in the slot of a template. The version of CHREST used in this simulation had 100,000 chunks, and is representative of a chess master.

 Insert Fig. A1 about here

At 30 ms, no chunk has been encoded in STM. At 240 ms, the program has recognized a medium-sized chunk, which happens to be a template. It is a fairly common black castling constellation. However, at this point, CHREST did not have the opportunity to encode anything in the template slots and used only the core of the template. Note that the black bishop, the black knight, and one of the black pawns are incorrectly located. This is due to the fact that CHREST enables some fuzziness in the way patterns are matched, the information used to sort to a given node not being

necessarily identical to the information stored at this node.¹¹ At 490 ms, a second chunk has been recognized, and a template slot has also been filled (black bishop on c6).

At 750 ms, a second piece has been encoded in the template slots, and a third chunk has been recognized. This new chunk does not add new information, as the three pieces were already encoded by the main template. This is typical of the way CHREST works, where there is always the possibility that the information held in different chunks overlaps. At 850 ms and 960 ms, no new information has been added. At 1060 ms, the only progress is that the white pawn on c5 has been encoded in one of the template slots. This illustrates that the program has a fair amount of redundancy, as this pawn is now encoded both in a chunk and in a template.

At 1350 ms, a larger chunk on the king's side has been recognized, which correctly encodes the location of the black pawn on h6. There is now uncertainty as to whether CHREST would now replace the pawn on h7 or h6, and this type of uncertainty with lateral pawns is typical of human behaviour (De Groot & Gobet, 1996; Jongman, 1968). Given that visual STM is limited to three items, this new chunk has dislodged the small chunk on the black queen's side. The black queen has now been encoded as well in the template. At 1490 ms, the white knight on e5 has been encoded in the template, and at 1610 ms, the same happens to the black knight on f4. It is interesting to note that, just like most of the masters studied by de Groot

¹¹ Technically, this is due to the way the mechanisms of discrimination (construction of the network) and familiarisation (building of the information held at a given node) work together. The possible discrepancy between the information used to reach a node and the information stored at this node offers an important means of simulating errors in chess (De Groot & Gobet, 1996; Gobet & Simon, 2000) and in verbal learning (Feigenbaum & Simon, 1984).

and Gobet (1996), CHREST memorised the four perceptually salient pieces in the centre of the board (Pc5, Nd6, Ne5, and nf4).

At 1740 ms and 1850 ms, CHREST has encoded that the squares e8 and d8 are empty. This is not particularly useful for a recall task, but this is clearly something that chess masters occasionally do (De Groot & Gobet, 1996), as empty squares are important strategically (Holding, 1985; Reynolds, 1982; Tikhomirov & Poznyanskaya, 1966). At 2100 ms, the program shows the same state of affairs.

We can now observe in a fast forward mode what happens during the last three seconds. At 3110 ms, the program has recognized a medium-sized chunk on the white king's side, which also happens to be a template. The template has enabled the encoding of three white pieces. The situation remains unchanged at 4000 ms and at 4840 ms. Thus, at the end of the presentation of the position, the program would replace 21 pieces correctly out of a total of 24 pieces. Three pieces are missing (all on the "a" column), and three placements would be counted as errors of commission (black bishop on e7, black knight on f6, and black pawn on h7). (While hesitating about the placement of a pawn or a piece, such as the black pawn h7/h6 in our example, humans either go for one location or replace both of them.) This amount of recall is fairly consistent with what has been observed in the literature with strong masters (Gobet et al., 2004).

What this example illustrates is that CHREST incrementally constructs a representation of the position in memory, and that the access to templates seriously boosts its memory. In particular, the recall performance would be fairly low without the possibility of encoding information in templates (that is, without the pieces on greyed squares). This is consistent with what has been observed with human masters

when the presentation time ranges from one second to 60 seconds (Gobet & Simon, 2000).

The astute reader may object that it is unclear how the program would see the checkmate threat (Qg5xg2; the queen is protected by the black bishop on c6 and the black knight on f4) within five seconds, given that the information on the white king's side has been perceived fairly late, and no single chunk encodes both the black queen and the white king. Interestingly, out of the four human masters whose behaviour is discussed in De Groot and Gobet (1996) with respect to this position, *only one saw this threat*. Indeed, whereas masters reliably see threats when the attacking and attacked pieces are either close together or when one of the pieces is perceptually salient, as noted by Jongman (1968), the data of De Groot and Gobet (1996) clearly show that even "obvious" threats are often overlooked when these characteristics are lacking. Thus, the detection of threats, which of course plays a key role in understanding the meaning of a position, does not always operate automatically. In this respect, the trace produced by CHREST is consistent with the empirical data.

Figure captions

Fig. 1. Illustration of how perceptual chunks can implement the notion of a production. Patterns on the board (the circled groups of pieces) might elicit perceptual chunks in long-term memory, the condition part of the production. Some of these chunks (in this case the one elicited by the pattern on the lower right hand side of the board) might suggest possible moves, the action part of the production (here, the white bishop retreating to the square “f1” to parry the checkmate threat on “g2”). Productions operate unconsciously and intuitively and, with strong players, may lead to actions that are readily solutions to a problem.

Fig. 2. Overview of the key perceptual and memory mechanisms embodied in the template theory. A simulated eye selects patterns on the external board. These patterns are sorted through a discrimination net, which enables access to a chunk (node) in long-term memory. Chunks give access to diverse types of information in addition to the location of pieces (depicted in the Fig.), including, in the case of chess, what kinds of move should be played or what plan should be followed (additional information is not shown in the Fig.). Information accessed in long-term memory is then placed in short-term memory, which consists of a queue of chunks and a pictorial short-term memory, where visuo-spatial information can be unpacked.

Fig. 3. Illustration of the mechanism of template formation (Panel a, discrimination net; Panel b, representation of the piece location on the chessboard). If a given type of information recurs often below a node in the discrimination network, a slot can be created at this node, specifying both the variable and the values that this variable can have. For example, given that a white pawn is used in

three branches below the node depicted in grey, a slot can be created for “white pawn”, and the possible values for the squares on which the pawn can be located are “f2,” “e4,” and “d4.” Similarly, a slot can be created for the square “e4,” which can have the value “black bishop” and “white pawn.” Chunks possessing slots are called templates.

Fig. 4. Sample of templates created by CHREST for chess.

Fig. 5. Pattern of eye movements for a chess master (top) and for a CHREST simulation (bottom). The semantically important squares are displayed in grey. (After De Groot & Gobet, 1996. Reproduced with permission of the copyright holder.)

Fig. A1. The time course dynamics of STM during the 5-s viewing of a chess position, according to CHREST.

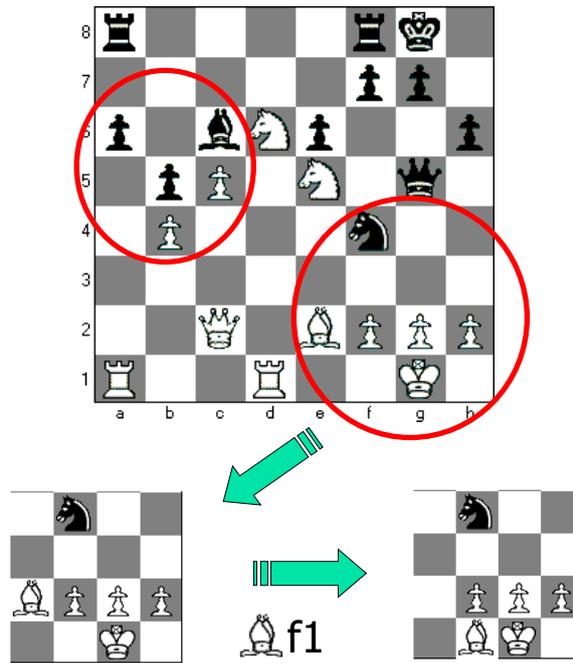


Fig. 1

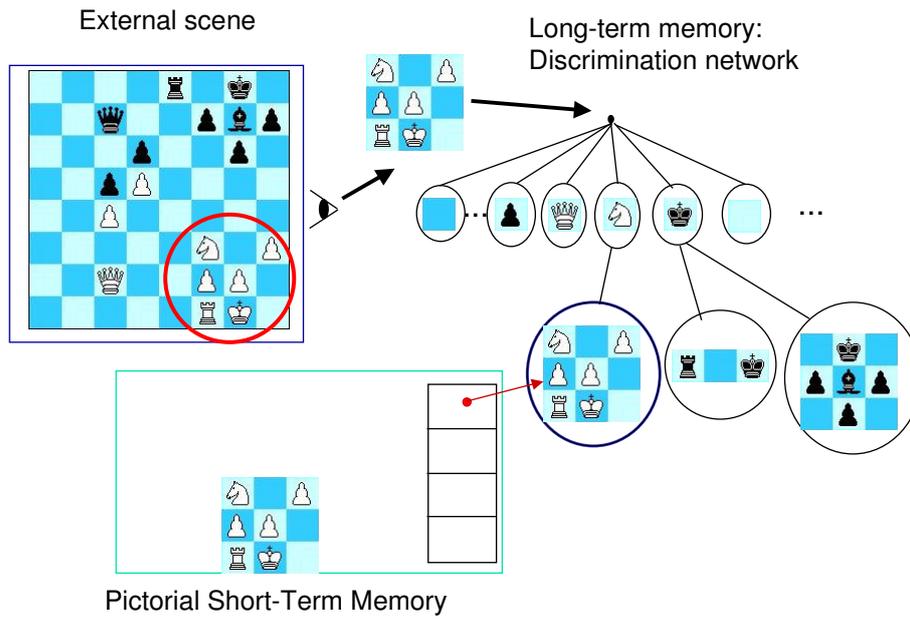
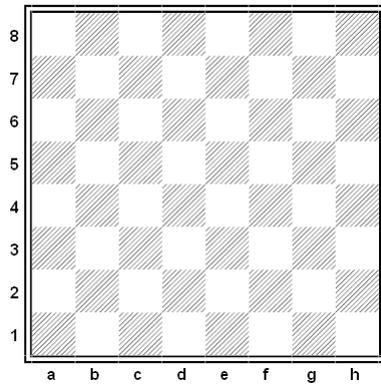


Fig. 2

30 ms

STM empty

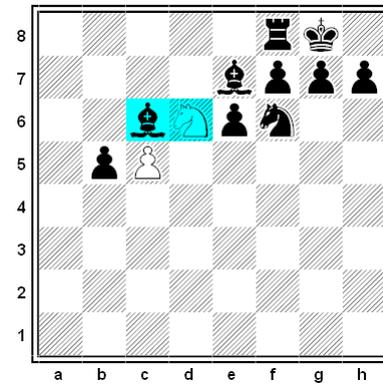


750 ms

#T pe6 pf7 pg7 ph7 kg8 be7 nf6 rf8

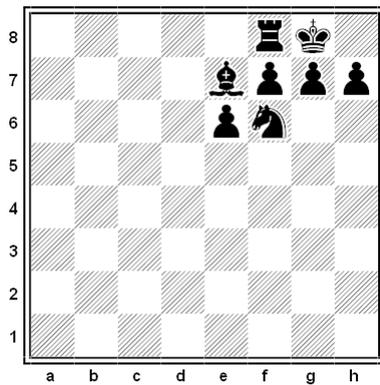
#C pe6 pf7 bc6

#C Pc5 pb5



240 ms

#T pe6 pf7 pg7 ph7 kg8 be7 nf6 rf8

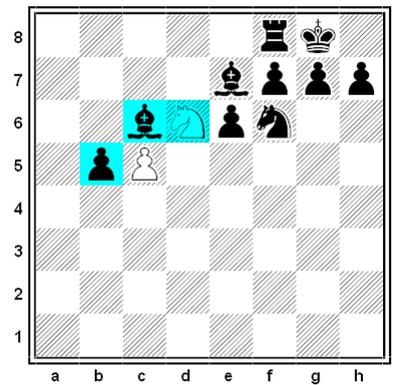


850 ms

#T pe6 pf7 pg7 ph7 kg8 be7 nf6 rf8

#C pe6 pf7 bc6

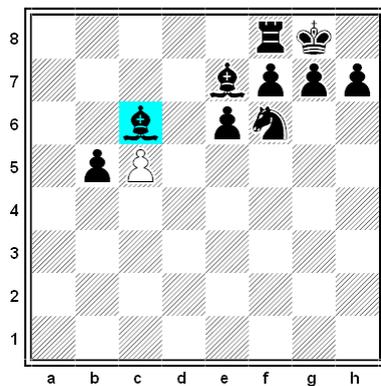
#C Pc5 pb5



490 ms

#T pe6 pf7 pg7 ph7 kg8 be7 nf6 rf8

#C Pc5 pb5

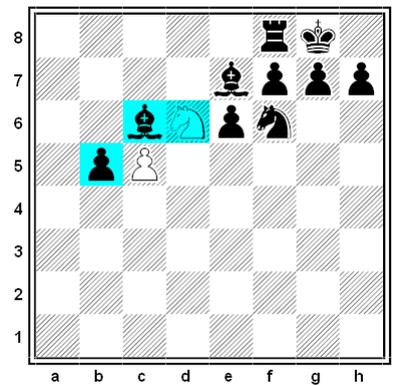


960 ms

#T pe6 pf7 pg7 ph7 kg8 be7 nf6 rf8

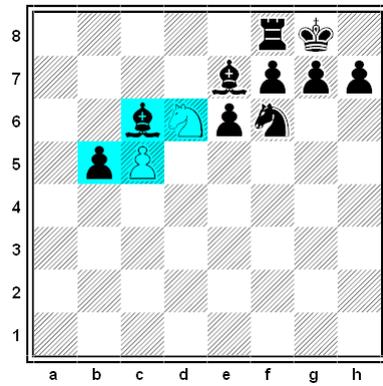
#C pe6 pf7 bc6

#C Pc5 pb5



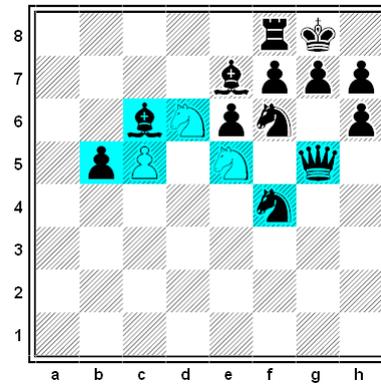
1060 ms

```
#T pe6 pf7 pg7 ph7 kg8 be7 nf6 rf8
#C pe6 pf7 bc6
#C Pc5 pb5
```



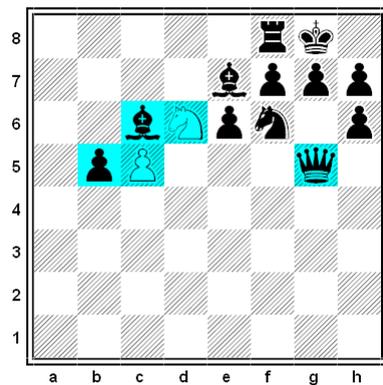
1610 ms

```
#T pe6 pf7 pg7 ph7 kg8 be7 nf6 rf8
#C pe6 pf7 pg7 ph6 kg8
#C pe6 pf7 bc6
```



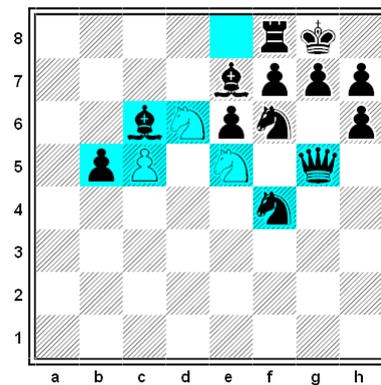
1350 ms

```
#T pe6 pf7 pg7 ph7 kg8 be7 nf6 rf8
#C pe6 pf7 pg7 ph6 kg8
#C pe6 pf7 bc6
```



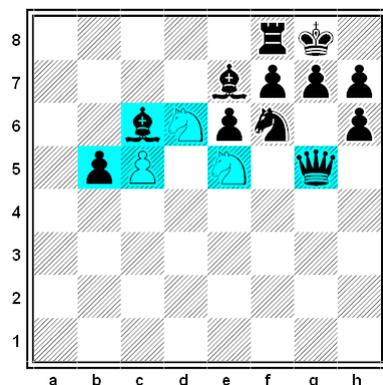
1740 ms

```
#T pe6 pf7 pg7 ph7 kg8 be7 nf6 rf8
#C pe6 pf7 pg7 ph6 kg8
#C pe6 pf7 bc6
```



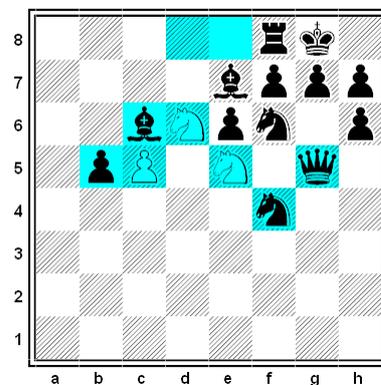
1490 ms

```
#T pe6 pf7 pg7 ph7 kg8 be7 nf6 rf8
#C pe6 pf7 pg7 ph6 kg8
#C pe6 pf7 bc6
```



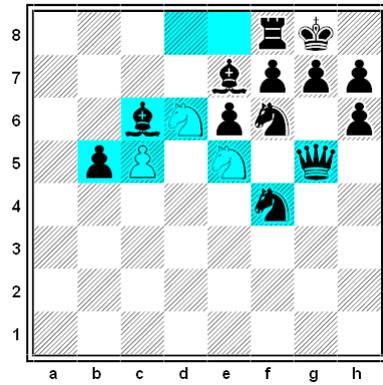
1850 ms

```
#T pe6 pf7 pg7 ph7 kg8 be7 nf6 rf8
#C pe6 pf7 pg7 ph6 kg8
#C pe6 pf7 bc6
```



2100 ms

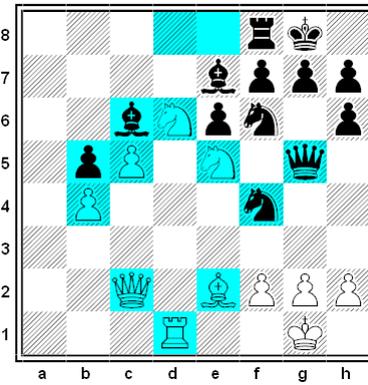
#T pe6 pf7 pg7 ph7 kg8 be7 nf6 rf8
 #C Pc5 pb5
 #C pe6 pf7 pg7 ph6 kg8



(...)

4000 ms

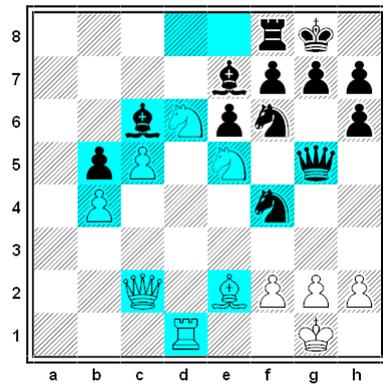
#T pe6 pf7 pg7 ph7 kg8 be7 nf6 rf8
 #C pe6 pf7 pg7 ph6 kg8
 #T Pf2 Pg2 Ph2 Kg1



(...)

3110 ms

#T pe6 pf7 pg7 ph7 kg8 be7 nf6 rf8
 #C pe6 pf7 pg7 ph6 kg8
 #T Pf2 Pg2 Ph2 Kg1



(...)

4840 ms

#T pe6 pf7 pg7 ph7 kg8 be7 nf6 rf8
 #C pe6 pf7 pg7 ph6 kg8
 #T Pf2 Pg2 Ph2 Kg1

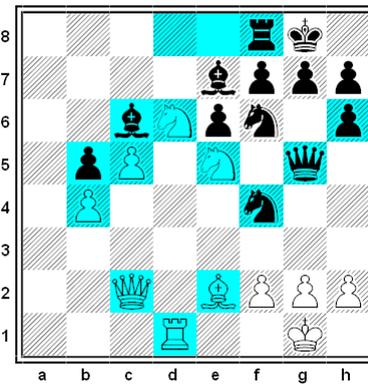


Figure A1