Running Head: INFLEXIBILITY OF EXPERTS

Inflexibility of Experts – Reality or Myth?

Quantifying the Einstellung Effect in Chess Masters

Merim Bilalić

Oxford University

Peter McLeod

Oxford University

and

Fernand Gobet

Brunel University

Correspondence Address:

Fernand Gobet

School of Social Sciences

Uxbridge

Middlesex, UB7 3PH, Great Britain

E-mail: fernand.gobet@brunel.ac.uk

Phone: +44 1895 265 484

Fax: +44 1895 237 573

Abstract

How does the knowledge of experts affect their behaviour in situations that require unusual methods of dealing? One possibility, loosely originating in research on creativity and skill acquisition, is that an increase in expertise can lead to inflexibility of thought due to automation of procedures. Yet another possibility, based on expertise research, is that experts' knowledge leads to flexibility of thought. We tested these two possibilities in a series of experiments using the Einstellung (set) effect paradigm. Chess players tried to solve problems that had both a familiar but non-optimal solution and a better but less familiar one. The more familiar solution induced the Einstellung (set) effect even in experts, preventing them from finding the optimal solution. The presence of the non-optimal solution reduced experts' problem solving ability was reduced to about that of players three standard deviations lower in skill level by the presence of the non-optimal solution. Inflexibility of thought induced by prior knowledge (i.e., the blocking effect of the familiar solution) was shown by experts but the more expert they were, the less prone they were to the effect. Inflexibility of experts is both reality and myth. But the greater the level of expertise, the more of a myth it becomes.

Key words: Flexibility, Expertise, Einstellung (set) effect, Fixation, Skill acquisition, Automatization, Creativity, Education, Chess

Inflexibility of Experts – Reality or Myth?

Quantifying the Einstellung Effect in Chess Masters

The knowledge base of the expert permits feats that seem incredible to the novice. An expert chess player, for example, can play several games simultaneously without sight of the boards. But, paradoxically, it has been argued that experts may fail on problems that novices solve. When a novel approach is required, the experts' knowledge can make them unable to adapt to the new task demands. Sternberg (1996) summarised this view of the inflexibility of experts: "…there are costs as well as benefits to expertise. One such cost is increased rigidity: The expert can become so entrenched in a point of view or a way of doing things that it becomes hard to see things differently." (p. 347).

In this paper we will explore the question of expert (in)flexibility using the Einstellung paradigm of Luchins (1942) with expert chess players. Given that the empirical evidence for expert (in)flexibility is rather sparse, we will first look at potential theoretical arguments for both possibilities – experts' flexibility and inflexibility. We will briefly review research in creativity and skill acquisition that is cited to predict the paradoxical possibility that experts, for all their knowledge, may become inflexible. The theoretical arguments for expert flexibility from research based on the natural study of experts will then be presented. We will then review empirical evidence suggesting that experts may indeed become inflexible in situations where a novel method or at least a modification of the existing method of dealing with the problem at hand is necessary. We will sketch out a possible theory for expert flexibility based on the previous empirical evidence and will finally present empirical evidence which provides new insight into the issue of expert (in)flexibility. Our main goal, however, will be to try to uncover the mechanisms behind experts' (in)ability to resist the Einstellung effect. This effect occurs in situations where people are unable to choose a less familiar, but optimal solution, rather than a familiar but non-optimal solution.

Creativity

When looking for possible reasons for expert inflexibility, researchers often point to work on creativity where a tension between knowledge (expertise) and inventive problem solving seems apparent (e.g., Frensch & Sternberg, 1989; Hesketh, 1997; Sternberg, 1996; Sternberg & Frensch, 1992; Zeitz, 1997). Although there is no universally accepted definition of creativity, most researchers agree that a creative product should be original and useful (see Sternberg, 1999). Given such characteristics, the creative product should go beyond previous knowledge and expertise, break links with the past and move away from stereotypical thinking. Implicit in this notion of creativity is that knowledge is necessary but not sufficient (for a different view see Hayes, 1989; Kulkarni & Simon, 1988; Weisberg, 2006).

The ambivalent role of prior knowledge in creativity can be seen in many different research traditions. Gestalt psychology distinguished between reproductive thinking as the application of previously acquired knowledge and productive thinking as the ability to go beyond past experience and produce something new (Wertheimer, 1959); James (1899/1939) talked about "skilled automata"; Koestler (1989) emphasized the necessity of avoiding habits; Amabile (1989) saw breaking out of an established pattern as the first step in the development of creative thinking (see also Lubart, 1994; Sternberg & Lubart, 1995). The following quote from Kuhn's famous work on 'The Structure of Scientific Revolutions' (1996/1962) summarizes the view that very knowledgeable people can be inflexible: "Almost always the men who achieve these fundamental inventions… have been either very young or very new to the field whose paradigm they change" (p. 90). These creative minds were more open to new ideas and could accept the break with previous theories because they were not too deeply immersed in the established thought patterns of their more experienced colleagues.

According to this view, knowledge is necessary but too much of it can be harmful for creativity – it is almost a necessary evil (see also Csikszentmihalyi, 1996; Simonton, 1999).

This view of creativity is not restricted to scientific research. As a popular 'creativity trainer' wrote: "Too much experience within a field may restrict creativity because you know so well how things *should be done* that you might be unable to escape to come up with new ideas" (DeBono, 1968, p. 228). Consequently, it is possible that experts, whose superior performance is the consequence of acquired knowledge, will run into difficulties when confronted with a situation that requires a novel approach going beyond their knowledge base.

Skill Acquisition

The other area of research that is cited as providing theoretical reasons for expert inflexibility is skill acquisition with its central concept of automatization. Automatization has been seen as central to skill acquisition ever since Bryan and Harter (1899) claimed that higher level habits cannot be acquired unless lower level habits had been automated. Shiffrin and Schneider (1977; Schneider and Shiffrin, 1977), for example, made direct use of automatization in their distinction between controlled and automatic processing. While controlled processing is effortful, usually slow and error prone, an automatic process is "a fast, parallel, fairly effortless process, not limited by … memory capacity, (and) is not under direct subject control" (Schneider, Dumais, & Shiffrin, 1984; p. 1). Although automatic processing has been thought to occur under consistent training conditions only, it has been shown that automatization can be developed even when the stimuli are mostly, but not completely, consistent (Schneider & Fisk, 1982).

A different approach was taken in Fitts' three-stage model of skill acquisition (Fitts, 1964; Fitts & Posner, 1967) where automatization develops gradually through practice. Nevertheless, performance in the final phase is fast, effortless, and hardly susceptible to voluntary control just like the automatic processing described by Schneider and Shiffrin (1977). Although Fitts' framework was designed to describe perceptual-motor skill acquisition, the situation with cognitive skill is assumed to be similar (see, for example, Rosenbaum, Carlson, & Gilmore, 2001). In Anderson's ACT-R framework (1983; Anderson & Lebiere, 1998) cognitive skill acquisition starts with a declarative phase, where the learner acquires general information and knowledge about the task, and finishes with a procedural phase in which the knowledge from the declarative phase has been turned into productions, compiled (combination of a series of simpler productions collapsed into a more complex production), and strengthened (productions that produce better results are preferred over the productions that do not). The end result is fast and accurate performance in familiar situations. The ACT-R framework does not directly feature the notion of automatization, but the strength of a production indirectly determines the degree of automatization (Anderson, 1992).

Unlike the stage models of skill acquisition just described, a number of models suppose a strategy shift in performance rather than increased efficiency of processes. In the Instance Theory of Automaticity (Logan, 1988) automaticity is a transition from a general algorithm used to solve problems to retrieval of past solutions from memory. At first, individuals rely on crude, slow and effortful strategies for obtaining the solution to the problem at hand. With each encounter of a stimulus, however, individuals store relevant traces/instances of the stimulus that can later be retrieved when they encounter the same stimulus/problem. The performance of a task is determined by two kinds of race. The first is between specific instances in memory and a general algorithm, while the second is between the instances themselves, the number of which grows with each encounter with the stimulus. Instance Theory traces automatization processes back to memory retrieval, that is when instances start winning the race against general algorithms. (For a different version of a similar memory based proposal, see Rickard, 1997; 2004; for a proposal about information reduction as a strategy shift, see Haider & Frensch, 1996; 1999a; 1999b.)

There are differences in the explanation of automatization in these theories of skill acquisition. Nevertheless, in all the frameworks, performance in the final phase of skill acquisition is fast, accurate, mostly effortless, seemingly beyond conscious control, in a word – automated. Automatization is generally considered advantageous because it saves time and cognitive resources that can be directed to other aspects of the task. But it is easy to see how this favourable aspect of skill can turn against the skilled individual. When the general features of a task trigger off an automated response but another, uncommon, response is required, experts may become victims of their automated task execution which, once initiated, is difficult to control.

Flexible Experts

In contrast, theoretical reasons to believe that experts will be flexible come from research into expertise. Expertise research suggests that in learning to adapt to the structure of the domain, experts increase their repertoire of methods for solving problems (de Groot, 1946/1978; Ericsson & Lehman, 1996; Ericsson et al., 1993; Gobet & Waters, 2003; Simon, 1969). Theories assuming that experts acquire their knowledge through chunking mechanisms (chunking theory, Chase & Simon, 1973; template theory, Gobet & Simon, 1996a) explicitly propose that experts become more sensitive to the fine details of their environment by acquiring chunks of knowledge that become increasingly differentiated. (Computationally, these theories have been implemented using the formalism of discrimination networks - Gobet & Simon, 2000; Gobet et al., 2001). In parallel to acquiring more differentiated perceptual chunks, experts also acquire more differentiated actions. These can be linked by productions giving experts a more flexible range of potential behaviour than non-experts (Chase & Simon, 1973; Gobet & Jansen, 1994). When practice leads to sufficiently discriminated knowledge, chunks can act as templates (similar to schemata, Bartlett, 1932; Minsky, 1977) and variable information can be stored rapidly into the template slots (Gobet, 2000; Gobet & Simon, 1996a; 2000).

In the study of chess experts it has been shown that acquiring perceptual chunks, productions and templates increases experts' look-ahead ability, and therefore flexibility, in several ways (Gobet, 1997; Campitelli & Gobet, 2004). First, they can pick up finer details than weaker players, which allow them to be more sensitive to minimal contextual details. Second, the ability to group information using chunks and templates enables them to compensate for basic limitations in memory capacity and processing speed. Third, templates, which enable information to be rapidly encoded in LTM, limit the impact of STM memory decay, allow search at a more abstract level than moves, and allow rapid change in the information used to describe the chess position. As a consequence, the more expertise someone has, the harder it should be to 'fixate' them into a stereotypical pattern of behaviour.

The possibilities expressed in the previous paragraphs, that experts might display either flexibility or inflexibility, are based on theoretical implications of research into creativity, skill acquisition, and expertise. As we will shortly see, although the topic of expert (in)flexibility has frequently been discussed in the scientific literature (e.g., Ericsson, 1998, 2003; Feltovich, Spiro, & Coulson, 1997; Hesketh, 1997; Krems, 1995; Sternberg, 1996; Zeitz, 1997), the empirical evidence for either possibility is sparse and unconvincing.

(In)flexible Experts – empirical evidence

We can define flexibility of thought as the ability to adapt to problems where it is necessary either to use new methods, techniques, knowledge or information, or to modify the existing method of dealing with the problem. In this sense, inflexibility of thought induced by prior experience has been demonstrated experimentally in two situations. In the first, participants can solve the problem with their existing knowledge and do so failing to notice that a superior, but less familiar, solution is available. In the second, a change in the task requires a novel procedure to be generated but experts have difficulty doing this because they cannot suppress the procedures they already possess.

A classic example of the first sort of harmful influence of prior experience on problem solution is the Einstellung (set) effect demonstrated by Luchins (1942) with the water-jug problem. Participants were given five introductory problems that could be solved using the same method. These were termed 'Einstellung problems' because the common solution method used for all of them was expected to induce a mental set for solving similar problems. The next two problems, termed 'critical', could be solved using the same method but a shorter and simpler solution was also possible. Over 80% of the participants failed to notice the shorter method and continued to use the longer method they had used successfully in the introductory problems. Luchins then presented an 'extinction' problem. This could not be solved using the original method but could be solved with the shorter method. Sixty four percent of the participants failed to solve the extinction (1-solution) problem. In comparison, only 5% of a control group who started with the critical (2-solution) problems and did not experience the introductory problems failed to solve the extinction problem (Luchins & Luchins, 1959). Thus, rather than improving their performance, the experimental group's additional experience of the general problem situation blinded them to a simple solution which was found by almost everyone who had not had the extra experience. Luchins' striking demonstration that experience can induce inflexibility of thought has been successfully repeated many times in a variety of formats (e.g., Atwood & Polson, 1976; Chen & Mo, 2004; Delaney, Ericsson, & Knowles, 2004; Lippman, 1994; Lovett & Anderson, 1996; McKelvie, 1990; Woltz, Gardner, & Bell, 2000).

The second example of inflexibility has been demonstrated in a number of studies. Frensch and Sternberg (1989) modified the game of bridge and let experts and novices play against a computer with new rules in place. Although experts outperformed novices in the modified game (see Frensch & Sternberg, 1991), experts had more trouble with the adaptation to deep structural changes than the novices (who were more impaired by perceptual changes). Supposedly, the automatization of the experts' existing procedures made it difficult for them to generate or use new ones. Similarly, Hecht and Proffitt (1995) showed that waitresses and bartenders were less accurate in judging the angle of the surface of water when its container was tilted than people with less experience of serving drinks (the water level task; Piaget & Inhelder, 1948/1956). They suggested that this resulted from the experts' inability to stop using their usual problem representation. This representation is what makes them good at what they do – bringing drinks to customers without spilling them. When confronted with a problem that required them to change it and imagine spilling drinks, they had difficulty doing so. Wiley (1998) showed that knowledge can act as a mental set in problem solving. Participants with more baseball knowledge performed worse in a version of Mednick's (1962) Remote Association Test (RAT) on misleading items when their baseball knowledge suggested inappropriate solutions. The more they knew about baseball the more they were fixated in their search for appropriate associations.

Although the studies cited above show that experience can lead to inflexibility, they do not justify extrapolation to the general conclusion that experts will be inflexible. The participants in the water-jug problem can hardly be characterized as 'experts' after solving the introductory problems. The other studies tested the experts' skill in a different domain from that in which they acquired their knowledge. Changing the rules of a game, as in Frensch and Sternberg's experiment, creates a new game. Although this game might have some similarities with the previous one, it is a new game and bridge experts are not experts in that game. The waitresses and bartenders in Hecht and Proffitt's experiment were experts in serving people drinks, not in the water level task. They do posses more experience with liquids in glasses but that does not make them experts in the judgement required in the water level task.[1] Wiley showed that inappropriate knowledge can lead to fixation in a creativity test. However, people who possess more knowledge of baseball might be experts in a quiz on baseball but they are not experts in Mednick's RAT. Their knowledge of baseball is not relevant in the RAT, especially on trials that were intentionally created to produce an inappropriate association. In all of these studies, inflexibility was inferred from a difficulty in moving to a domain related to, but different from, that in which the knowledge that made them an expert had been acquired.

In contrast, Einstellung tasks seem to offer a good chance of demonstrating the inability of experts to adapt to new situations. Other tasks involving negative transfer between training and test situations are often easy to spot and consequently to correct by the person experiencing them because their performance suffers (Singley & Anderson, 1989). Einstellung phenomena, in contrast, are notoriously difficult to notice. In the critical, 2-solution problem of an Einstellung task an inefficient method of dealing with the problem is transferred but the method produces a solution. So the person experiencing the effect does not realise that there is any need to look for a better one.

The Einstellung Effect in Chess

Using the classical water-jug design, Saariluoma (1990; Experiments 2 and 3) showed that chess players are prone to fixation. He first gave players four Einstellung problems that could be solved using smothered mate, a well-known motif in chess. After the introductory problems, a critical problem was presented which contained two shorter solutions although the smothered mate could still be used. The critical (2-solution) problem was followed by another set of Einstellung problems, this time with one of the shorter solutions for half the players while the other half received problems that were solvable using the other shorter solution. Then both groups again encountered the 2-solution problem. Ten of the 12 players failed to notice the shorter solutions in the 2-solution problem the first time it was presented. However, when they again encountered the 2-solution problem after practicing problems which used one of the shorter solutions, both groups used the shorter solution they had been practicing. So the players were fixated twice –when the 2-solution problem appeared for the first time (they failed to find either of the shorter solutions) and when the 2-solution problem reappeared (they found the shorter solution but only the one they had been practicing – the other shorter solution remained unnoticed).

In Experiment 3 Saariluoma (1990) showed two matched groups of six players two sets of problems. The first group had to solve problems where an optimal solution existed but there was also a more familiar[2] but non-optimal one. Only 11% of the solutions chosen by this group were the optimal one. Most of their choices (67%) were the familiar but non-optimal solution. The other group solved similar problems in which the gist of the problems stayed the same but the more familiar solution was disabled by a change of the position of one piece. In these problems there was only one solution – the optimal one from the other group's problems. Now 69% of players found the optimal, less familiar solution. Saariluoma concluded that it was not the difficulty of the optimal solution that prevented players finding it but the familiarity of the other solution that blocked the optimal solution.

Theories of expert (in)flexibility

Saariluoma demonstrated that, despite their extensive knowledge, chess experts can fail to find an optimal solution. The familiar solution may be so salient that it blocks the optimal one. If we assume that the knowledge players acquire with experience is represented as chunks and templates (Gobet & Simon, 1996a), then it is possible to explain the salience of the familiar solution. Knowledge structures are represented as nodes in LTM that are also connected with possible actions or information (e.g., in chess, evaluation of the position or possible plans). Upon seeing a position, more integrated knowledge that is better indexed by more and larger chunks will be retrieved more easily. This could explain why players could see the well-known solution before any other. The optimal solution is less well indexed in LTM, due to its lower frequency of occurrence, and is more slowly retrieved. That is, fewer occurrences of the optimal solution in the chess environment mean that the key pattern has not been stored with a large variety of chunks denoting diverse contexts, which would enable a redundant and robust encoding of the information, as is the case with the familiar solution.

This theory, developed in the tradition of expertise research, may be contrasted with two theories from the skill acquisition tradition we have mentioned earlier. With ACT-R, once players encounter a familiar problem the strongest production connected with it will become activated (Anderson & Lebiere, 1998). With the instance theory, one of the many instances stored for previously encountered smothered mates will win the race against other instances and any general algorithm (Logan, 1988). In all three theories, this process of retrieving the familiar solution is supposed to be quick, effortless, and difficult to avoid. Given that the familiar solution is good enough and solves the problem at hand, experts are satisfied with it and usually spend little time looking for a better solution (as in typical Einstellung phenomena). Consequently, they choose the familiar, good-enough solution and overlook the optimal one. It is thus possible that experts are inflexible, that is, prone to choosing a familiar over the optimal solution. A plausible explanation of the general Einstellung phenomenon can be constructed within the framework of existing theories of expertise and skill acquisition.

Saariluoma (1990), however, reported that the only exceptions to the fixated group of players were “the one master who was not fixated the first time, and the grandmaster, who saw both alternatives” (p. 45). It seems that more skilful participants may be able to resist becoming fixated by the familiar but non-optimal solution. Given the current empirical evidence, it is not entirely clear how they were able to resist choosing the familiar solution.

One possibility is that with increased skill level players are more likely to immediately activate the optimal solution, although one needs to explain how the optimal solution is selected in spite of its less frequent exposure compared to the familiar solution.

Another possibility is that although both more and less skilful experts find the familiar solution first, in those few remaining moments as they double check the solution (which chess players usually do) more skilful experts notice the optimal solution while less skilful experts fail to do so. Yet another option appeals to motivational factors – more skilful experts simply look longer than less skilful colleagues. Consequently, more skilful experts are more likely to find the optimal solution.

We explored these possibilities with a modification of Saariluoma's design. Saariluoma (1990) showed that chess players can be fixated and show signs of inflexibility, just like the participants in the water-jug task. However, his goal was to show the existence of apperception (for a detailed explanation of this term, see Saariluoma, 1995), and not to examine the flexibility of experts. Consequently, his experiments are not ideal for examining this issue. First, there was no extinction (1-solution) task. Without an extinction task, it is not possible to know whether players who failed to find the shorter solutions failed because they were distracted by the more familiar smothered mate theme or because they did not know the motif featured in the shorter line. Second, the instruction was to find a checkmate as fast as possible. This instruction facilitated the Einstellung effect because participants had no time to consider alternatives after the first solution that they noticed. Third, Saariluoma's players were relatively weak. Only three had a USCF[3] rating above 2100.

These shortcomings and inconsistent findings preclude us from drawing firm conclusions about experts' behaviour in Einstellung tasks. We do not know how experts tackle these problems, whether they are able to resist the temptation of a well-known solution over an optimal one, why they succeed or fail to notice optimal solutions, and how these issues depend on their skill or/and on the problem difficulty. We will show that answers to these questions provide insights into the mechanisms behind the Einstellung phenomena that illustrate the role of prior knowledge on performance and the nature of expertise, and relate to issues education and creativity.

Experiment 1

To examine whether more expert chess players show less flexibility of thought than less expert players (that is, are less likely to consider alternatives when they have already found a good solution) we used Saariluoma's idea of the smothered mate theme producing the Einstellung effect in a design based on the water-jug problem. We did not use introductory problems as skilled chess players are already highly familiar with the smothered mate theme, so there is no need to induce this method of solving the problem. We included an extinction task (1-solution problem) to check whether players who failed to find the shorter solution in the critical task (2-solution problem) did indeed possess knowledge of the solution. The 2-solution problem was presented to players with a range of expertise from 3 to 5 standard deviations (SDs) above the mean level of all graded chess players. By the standards of most experiments on expertise all these players would be considered to be 'expert'. Finally, a group of less skilful players (between the mean level and 2 SDs above) was shown the 1-solution problem only, in order to establish the difficulty of the motif featured in the shorter solution of the 2-solution problem. This allows the Einstellung effect (i.e., the ability of the familiar solution to block access to the better one) to be quantified.

Method

Participants. There were two groups of players. The first (the experts) were given both the 2-solution (critical) and 1-solution (extinction) problem. The second group (less skilful players) were given only the 1-solution problem. Players were recruited from local chess clubs in Oxford, Tuzla, and Sarajevo as well as during the Bosnian team championship in Neum 2003. For details of the players and their Elo ratings[4] see Table 1.

Insert Table 1 about here

Stimuli. The two problems are shown in Figure 1. The first solution of the 2-solution (critical) problem was the well-known theme of smothered mate that solves the problem (gives checkmate) in five moves. (Smothered mate is an elegant and easily remembered motif, taught to all young chess players, where the strongest piece, the Queen, is sacrificed in order to draw one of the opponent's pieces onto a square that blocks the escape square for the King). The shorter, but less familiar solution, solves the problem in three moves. In the 1-solution (extinction) problem only the shorter solution was possible. Although the problems look identical, a change in the location of one piece (Black's white squared Bishop) means that smothered mate is not possible in the 1-solution problem.

Insert Figure 1 about here

Design and procedure. Participants in the first group (the experts) were told that they should look for the shortest way to win[5] in the 2-solution problem and that they would have unlimited time to find the solution. They were asked to think aloud as they tried to solve the problem (the procedure described in Ericsson & Simon, 1993, was applied). The problems were shown on a 15" screen laptop computer and the players' spoken commentary tape-recorded. Those who failed to find the shorter solution were then shown the 2-solution problem. The same instructions were given to the participants in the second group (the less skilful players) who were shown the 1-solution problem only. All participants were tested individually in a quiet room.

Results

There was a clear effect of skill level in the experts of Group 1 with an increasing proportion of players failing to find the shorter solution in the 2-solution problem as their skill level decreased (see Table 2). To quantify the effect of the players' skill level, we performed a logistic regression with results on the 2-solution problem as the dependent variable and the Elo rating of the players as the predictor variable. Rating had a significant effect ($B = .013$, $S.E. = .005$, $Z = 8.4$, $p = .004$) such that a decrease of 100 rating points reduced the odds of finding the shorter solution by a factor of 3.67. In other words, the probability of finding the shorter solution in the 2-solution problem for an average Grand Master (Elo 2550), International Master (Elo 2450), Master (Elo 2350), and Candidate Master (Elo 2150), was .64, .32, .12 and .01 respectively.

Insert Table 2 about here

There was no evidence of the Einstellung effect (that is, the presence of the well-known smothered mate solution preventing the discovery of the shorter one) with the strongest players. The Grand Masters all found the optimal solution in the 2-solution problem. However, there was clear evidence of the Einstellung effect with the 'ordinary' experts (3 SDs above average skill level).[6] All the players in Group 1 were capable of finding the optimal solution as they all found it in the 1-solution problem (see Table 2). But only two of the Masters and none of the Candidate Masters found the optimal solution in the 2-solution problem.[7]

Given that the only difference between the 2-solution and 1-solution problem was the presence/absence of the familiar solution, the Einstellung effect can be quantified by comparing the results of the players in the first group (the experts) in the 2-solution problem with the results of players in the second group (the weaker players) in the 1-solution problem. The performance of the International Masters (5 SDs above average) on the 2-solution problem was comparable to that of Class A players (2 SDs above average) on the 1-solution problem - 50% and 63% respectively. The performance of the Masters (4 SDs above average) was comparable to that of the Class B players (1 SD above average) - 18% and 13% respectively. The performance of the Candidate Masters (3SDs above the average skill level) (0%) was the same as the performance of the Class C players (average skill level) (0%). The presence of a familiar solution reduced the problem solving abilities of the experts to about that of players 3 SDs lower in skill.

The verbal protocols showed that the well-known solution (smothered mate) was the first thing all players noticed when presented with the 2-solution problem. Having spotted it, all players looked further to see whether there was a shorter solution. The Grand Masters (who all found the shorter solution) spent a median of 7 s before finding it. The International Masters and Masters who found the shorter solution spent a median of 9 and 10 s respectively. The players who did not find the shorter solution spent similar amounts of time looking for it (9, 8, and 8 s for International Masters, Masters, and Candidate Masters respectively) before declaring that the smothered mate was the solution.

The power of the Einstellung effect on the ordinary experts is illustrated by the inaccurate predictions of 12 experts (*M* Elo 2231, *SD* = 150) who were shown both solutions in the 2-solution problem and asked to guess what percentage of the players of different strengths would find the optimal solution. They predicted that 86% of International Masters, 74% of Masters, and 59% of Candidate Masters would find the shorter solution in the 2-

solution problem. The actual results were 50%, 18%, and 0%. One of them commented: "You'll have to find a harder problem than this. No Master would miss [the shorter solution] in this position."

Discussion

By conventional standards all the players in Group 1 were experts but there were differences in flexibility between them. The 'super' experts (more than 5 SDs above average skill level) were flexible. They first saw the good solution, just like the ordinary experts, but this did not prevent them finding a better one. In contrast, the 'ordinary' experts (3 SDs above average skill level) were inflexible. When they could see a good solution they were unable to find a better one although they saw it when the good solution was no longer available. We have demonstrated for the first time an expertise effect in flexibility in problem solving – the more knowledgeable the experts were, the less likely they were to be trapped by the immediate appeal of the familiar solution. More expertise provides more resources to resist the temptation to choose a familiar but non-optimal solution.

The Einstellung effect was remarkably powerful. It reduced the performance of ordinary experts to that of average players. The gap between ordinary experts and average players (about three SDs) is a dramatic gulf in skill. Normally, average players have no chance of beating ordinary experts. Yet, the Einstellung effect reduced the performance of ordinary experts to that of average players. The hidden power of the Einstellung effect was demonstrated by the inaccurate metacognition of another group of experts who seriously underestimated the effect of the presence of the familiar solution. Why does the Einstellung effect have such a dramatic effect on the performance of experts?

The protocol analyses showed that all players found the familiar solution first. Although the familiar solution was good enough, that is it solved the problem, all players looked for a shorter/better one. This behavior is probably a consequence of several factors:

the general nature of chess players who normally double check their solutions, the instruction which explicitly asked for the shortest solution, and the experimental setting where players are probably more suspicious. The players, nevertheless, did not spend a lot of time looking for a shorter solution. This illustrates the main danger of the Einstellung effect – there is no feedback available on the inappropriateness of the well-known solution. Consequently, players do not spend too much time pondering other solutions. In those few moments of reflection, only the most skilful experts managed to find the optimal solution.

It seems, therefore, that the optimal solution was indexed less strongly in LTM than the familiar solution, probably due to its lower frequency of occurrence, and was more slowly retrieved in both super and ordinary experts. However, the optimal solution was still not retrieved in those few seconds in ordinary experts. In super experts, however, the production for the optimal solution was indexed well enough (Gobet & Simon, 1996a) to fire in the few seconds they spent looking for a better solution. The other two theories can also explain why the familiar solution is retrieved first. To explain with only the best players found the optimal solution, on has to assume that the production connected with the optimal situation was stronger in super experts that ordinary experts (Anderson & Lebiere, 1998). Similarly, super experts had more instances stored for the optimal solution than ordinary experts (Logan, 1988). Assuming that players do not have to execute the familiar solution the very moment they spot it and can reflect on their first solution, as typically happens in chess, this two-stage model would provide an explanation of why super experts were nevertheless able to find the optimal solution in those few seconds when they looked further at the problem.

The other two possibilities mentioned in the introduction are precluded by the data. There does not seem to be an interaction between solutions and expertise in the sense that the optimal solution takes the place of the more familiar solution as expertise increases. Also, the finding that all players tried to find a shorter solution after spotting the smothered mate first, and that the players who found it spent no more time than the players who did not, rules out the motivational possibility that the players who found the shorter solution simply looked for longer. Although the results of the first experiment were clear-cut, the familiar solution featured one of the most memorable themes in chess (smothered mate) and that may have influenced the problem solving behavior of experts. In addition, the choice between the two critical solutions occurs not in the first move of white, which is the same in both solutions, but in the second move of white. This may have made the recognition of the key patterns harder, as this must be done in the position anticipated in the mind's eye rather than that on the external board. Consequently, we wanted to see whether the findings of Experiment 1 generalize to tactical motifs that are more likely to occur in a normal game, where the candidate solutions diverge already in the first move, and with the more natural instruction to look for the best move available.

Experiment 2

Saariluoma's (1990) third experiment showed that the Einstellung effect can also be found in a wider range of tactical problems in chess. However, he did not use an extinction task (a 1-solution problem) so we do not know whether the players who failed to find the optimal solution failed because of the blocking influence of the familiar solution or because of lack of knowledge of the motif required for the optimal solution. We took three problems from Saariluoma's sample and two more that we created ourselves. In each case there was a familiar continuation that would lead to a clear advantage but also a less familiar continuation that would lead to a greater advantage. We also created analogous positions for the 1-solution problems. These could only be solved using the optimal solution from the 2-solution problems, the familiar solution having been disabled by replacing or moving some of the pieces. As in the first experiment a group of less skilful players were presented with the 1-solution problems to assess the difficulty of the optimal solutions in the 2-solution problems.

However, this time we asked players to find 'the best move', rather than to look for the shortest way to win as in the previous experiment. This is a more natural instruction that resembles the situation chess players face during play.

Method

Participants. The first group (the experts) were presented with both 2-solution (critical) and 1-solution (extinction) problems. (See Table 1 for details of the players.) These players were different from those in the previous experiment. The second group were the same 24 weaker players who tried to solve the 1-solution problem in Experiment 1. The players of both groups were recruited from local clubs in Oxford, Tuzla, and Sarajevo, as well as during the Bosnian team championship in Neum 2003.

Stimuli. There were five problems, each in a 2-solution (critical) and 1-solution (extinction) version. One of the problems (number 5) was considerably easier than the others. One of the problem pairs is shown in Figure 2. The other positions and the more and less familiar solutions can be found in Appendix A.

Insert Figure 2 about here

Design and procedure. The experts in Group 1 were instructed to look for the best move in the 2-solution problems. They were asked to think aloud (Ericsson & Simon, 1993) and told that thinking time was limited to five minutes. The presentation order of the problems was randomized across players with the exception that the easy position 5 was always presented last (in order to avoid suspicion). Participants who failed to find the optimal solution in the 2-solution version were given the corresponding 1-solution version of the same problems after they finished the five 2-solution problems. The second group (the weaker players) had the same instructions. We wanted to see if we could obtain an

Einstellung effect in weaker players with an easy problem. So they were given the 2-solution version of problem 5 first and then its 1-solution version if they failed to find the optimal solution. Then they were given the other 1-solution problems in random order. (They were not given the 2-solution positions for the first four problems.)[8] All players were tested individually in a quiet room.

Results

The percentages of players at each skill level in the first group (the experts) who found the optimal solutions to the first four 2-solution (critical) problems and the percentages of the less skilled players who found the solutions to the 1-solution (extinction) version of these problems are shown in Table 3. All participants in Group 1 who failed to find the optimal solution in any of the 2-solution problems found it in the 1-solution problem (not shown in Table 3). Hence, all players in the expert group possessed the necessary knowledge to solve the 2-solution problems. (The 2-solution version of problem 5 was solved by all players in Group 1 and the 1-solution version by all players in Group 2. So it did not distinguish between players of different skill levels in Group 1 and will not be discussed below. The ability of players in Group 2 to solve the 2-solution version of this problem did vary with skill level. We will discuss this later.)

Insert Table 3 about here

As Table 3 shows, the less skilful the experts were, the less likely they were to find the optimal solution in the first four 2-solution problems. To check whether problem, order of presentation, or interaction between the two had any effect on the results, we performed a logistic regression with repeated measures using the generalized estimating equation (GEE) method (Liang & Zeger, 1986) with the probability of finding the optimal solution as the dependent variable and rating, problem, order of presentation, and interaction between problem and order of presentation as predictors. Problem, order of presentation and their interaction had no effect on the results (all $Z < .5$, ns). Only rating had a significant effect ($B = .006$, $S.E. = .0016$, $Z = 3.72$, $p = .0002$). A decrease of 100 rating points decreased the odds of finding the optimal solution by a factor of 1.82. For example, the probability of an average International or Grand Master (Elo 2500) finding the optimal solution in the 2-solution problem in Figure 2 was .87, while the probabilities of an average Master (Elo 2300) and Candidate Master (Elo 2100) were .67 and .37, respectively.

The weaker players were relatively successful in solving the 1-solution problems showing that the solutions that the experts failed to find were not too difficult. The performance of the Class A players (2 SDs above average) with the 1-solution problem was comparable to that of the International and Grand Masters (5 SDs above average) with the 2-solution problem – 91% vs. 92%. The performance of the Class B players (1 SD above average) was comparable to that of the Masters (4 SDs above average) – 66% vs. 69%. The performance of the Class C players (average skill level) was comparable to that of the Candidate Masters (3 SDs above average) – 47% vs. 35%. Just as in Experiment 1, there was a powerful Einstellung effect. The presence of the obvious solution reduced the problem solving ability of the experts to that of players about 3 SDs below them in skill level.

However, the inflexibility of the ordinary experts (the Candidate Masters) should be kept in perspective. All the Candidate Masters found the optimal solution in the 2-solution version of problem 5. Thus the inflexibility of the ordinary experts interacted with the difficulty of the task. When the optimal solution was relatively hard to find (as in problems 1-4) they were subject to the Einstellung effect. But when the problem was easier (although still sufficiently difficult that less than half the class C players found the solution in the 2-solution problem) they were not fixated by the familiar solution. They now appeared flexible just like Grand Masters with the harder problems.

As in the previous experiment, the extent of the Einstellung effect was underestimated by the same group of twelve experts used Experiment 1. They estimated that 97% of International and Grand Masters, 89% of Masters, and 73% of Candidate Masters would find the optimal solutions in the four problems. The actual numbers were 92%, 69% and 35%.

The protocols showed that, unlike in the first experiment, some of the players found the optimal solution immediately without noticing the solution that was supposed to be more familiar. The percentage of players who immediately found the optimal solution for each problem decreased as their skill level went down (see Table 4). All positions showed a similar pattern of results except Problem 3 where none of the players found the optimal solution immediately. It can be argued that the more skilful players are more likely to find the optimal solutions not because they are more flexible but because they see them immediately. However, if we consider only those problems where the optimal solution was not found immediately, there was still a clear skill effect. In almost all instances (11 out of 12 – 92%) where the International/Grand Masters did not immediately find the optimal solution when solving problems 1, 2, and 4, they found it in the end. In contrast, in only 13 out of the 24 (54%) instances where the Masters did not immediately see the optimal solution, did they find it later. And in only 11 out of 32 (34%) occasions where the Candidate Masters did not immediately find the optimal solution, did they manage to find it. In addition, a *t*-test performed on Problem 3, where all players noticed the familiar solution first, showed that the players who found the optimal solution (*M* Elo = 2357) had a significantly higher rating than the players who failed to find it (*M* Elo = 2242), $t(34) = 2.09$, $p = .044$.

Insert Table 4 here

As in Experiment 1, the protocols revealed that all the players who chose the familiar non-optimal solutions had looked for a better one. The median time spent on looking for a better solution across the four problems for International/Grand Masters, Masters, and Candidate Masters who found the optimal solution (9.5, 12, and 20 s, respectively), were not significantly longer than those of the players from the corresponding category who did not find the optimal solution (9, 13, and 22 s).

Discussion

With a more natural context (the instruction to look for the best move) we replicated the finding of Experiment 1 – flexibility is positively associated with expertise. As before we found that the Einstellung effect was powerful (even very skilled chess players sometimes failed to find the optimal solutions in the presence of a more obvious but inferior solution) and that it was unexpected by other experts.

The problems in this experiment were easier for experts than the problem featuring the smothered mate theme in the first experiment (probably because the choice between the two solutions was already apparent by the first move whereas it did not appear until the second move in the first experiment). Ordinary experts found some of the optimal solutions in the 2-solution problems. Demonstration of the full Einstellung effect thus seems to depend on an interaction between the skill of the individual and the difficulty of the less obvious solution. With problems where the optimal but less obvious solution is easy to find (like Problem 5) the Einstellung effect could no longer be demonstrated with ordinary experts because they all found the optimal solution to the 2-solution problem. But with this problem the Einstellung effect could now be demonstrated with average players (Class C).

What was different in the more natural context was the apparent switch with some experts between the supposedly familiar and less familiar solutions. Unlike the previous experiment, where the familiar solution was without exception spotted first, in the current setting there were instances when the supposedly less familiar (but optimal) solution came to mind immediately. We can thus see that expert flexibility can have two forms: in some cases, it allows them to find the optimal solution immediately. In other cases, where the optimal solution is found only after the typical solution, some kind of search appears to be involved. Knowledge, however, seems to lie at the heart of both forms of flexibility. Even when strong players did not find the optimal solution right away, they presumably used their knowledge to guide search (Gobet, 1997; Gobet & Simon, 1996a). This assumption is supported by the short duration of the search.

Both experiments show that although experts can be trapped by the immediate appeal of a well-known solution to a problem, the more expertise players possess the more likely they are to find the optimal solution once they start to look further. However, both experiments made use of well-known tactical motifs. It might be argued that, since the less familiar solutions in the second experiment were more familiar to the strongest players, the results are an effect of familiarity rather than flexibility. Although this does not seem to be the case because the success rates for the problems that did not immediately elicit the optimal solution favoured more skilled players, in the third experiment we set up a situation where the discrepancy in familiarity between super experts and ordinary experts with the solution would be minimized.

Experiment 3

All positions in the second experiment were of a tactical nature (involving material sacrifice). Although tactical problems are not rare in chess, most of the problems that a player encounters are of a strategic nature (finding the right plan, a way of continuing the game). In Experiment 3, in consultation with an International Master who is highly familiar with the particular line of play we used, we constructed three problems where the solutions were strategic (see also Saariluoma & Hohlfeld, 1994).

In the first problem a definitive plan existed. In the second problem, which was different only in one crucial detail, the previous plan was impossible. Both problems are ‘theoretical positions’, meaning that the players who play the line know about the standard plan and about the crucial detail, and that the best plan applicable in the first problem is not a good option in the second. The critical problem was the third one. Here the key detail from the second problem, which prevented the plan of play in the first problem, remained. However, there were a few other slight differences that made the application of the plan from the first problem possible again. We wanted to see whether players familiar with the basic problem and the crucial detail in the second problem would be able to resist the fixation of not carrying out the plan in the third problem because of the presence of the crucial detail. Hence, we used players who play the line and who could easily solve the first two problems. If they were inflexible, they would discard the right plan in problem 3 because the position contained the critical feature. On the other hand, flexibility would mean that they would find the right plan despite the critical feature.

Insert Figure 3 here

Method

Participants. Twelve players were chosen who are specialists in the line of play from which the positions were taken (see Table 1). They were recruited from the Oxford University Chess Club, through personal contacts or during the Bosnian and British team championships 2003-4. None of the players participated in the previous two experiments.

Stimuli. The three problems were presented one after another. In the first, a well-known plan was applicable. In the second (in which the location of one piece was changed, the crucial feature) the plan was not applicable as it would lead to an inferior position. Both positions were familiar to the participants and they were expected to solve them quickly using their previous knowledge. The third and critical problem[9] featured the slight change but also several other changes (still keeping the gist of the position intact). The other changes made the application of the plan possible despite the presence of the critical feature. A brief analysis of the problems is presented in Appendix B.

Procedure. All players solved the first position correctly and were then shown the second position. All players solved the second position correctly and were then asked to solve the third and final position. For each position, they were given 10 minutes in which to find the best move during which time they were instructed to think aloud (Ericsson & Simon, 1993). All players were tested individually in a hotel room or the office of the first author.

Results and Discussion

The third problem proved to be very difficult. One of the five Grand Masters and all the other players failed to find the best solution. A Mann-Whitney *U* test using rating as the dependent variable showed that players who found the right solution (*Mdn* Elo = 2547) were significantly more skilful than players who did not (*Mdn* Elo = 2258); $U = 2$, $p = .017$.

The main difficulty in finding the correct solution to the third problem was the crucial feature. In the second problem, the standard solution was (correctly) rejected because of the feature. It seems that it is difficult to take the plan into consideration again in the next problem when the critical feature is still present. Despite the other changes that neutralize the feature and make the execution of the plan possible, the players were reluctant to consider it again. Most of the players mentioned the right plan as a possibility but discarded it because they evaluated it negatively. The best players, however, were also not satisfied with the evaluation of other plans and proceeded to examine the right plan in more detail. This further examination enabled them to realise the impact of the other changes, break the fixation produced by the critical feature, and integrate all the crucial parts of the problem to discover the right solution.

Although it is tempting to argue that the best players found the right solution not because they possessed necessary knowledge which less skilled chess players lacked but because they were more flexible, there are reasons why we should not reject the role of knowledge all together. The best players did evaluate the problem better than their weaker colleagues, leading them to reconsider previously abandoned plans. The evaluation of problem situations is heavily influenced by available knowledge (Gobet, 1997; Holding, 1985; Saariluoma, 1995). Knowledge is an essential part of flexible behaviour.

General discussion

Expert flexibility has been frequently discussed in the scientific literature (e.g., Ericsson, 1998, 2003; Feltovich, Spiro, & Coulson, 1997; Hesketh, 1997; Krems, 1995; Sternberg, 1996; Zeitz, 1997) but there have been few empirical investigations of this topic. Research in creativity and skill acquisition has been used to illustrate the possibility that expertise may and can lead to inflexibility – developing successful patterns of thought can block novel thoughts that may be more appropriate for the task (e.g., Frensch & Sternberg, 1989; Hesketh, 1997; Sternberg, 1996; Sternberg & Frensch, 1992; Zeitz, 1997). In contrast, the naturalistic study of experts suggests that the greater their level of expertise the more flexible and creative their thought patterns will be (Ericsson, 2003; Ericsson & Charness, 1994; Ericsson & Delaney, 1999; Richman et al., 1996).

The contrasting performance of our ‘ordinary’ experts (around 3 SDs above average performance) and our ‘super’ experts (more than 5 SDs above average performance) suggests that both options are possible depending on the expertise level and the problem difficulty. In the first experiment the ordinary experts all spotted a well-known solution. They looked for a better one but failed to find it. Similarly, the super experts all found the well-known solution first and then looked for a better one. In contrast to the ordinary experts, they all found it. The ordinary experts all found the better solution in the 1-solution problems where the good solution had been disabled. The same pattern was observed in the second experiment, in a more general and natural context. Although now, when the problems were more difficult, even the super-experts showed some inflexibility and when they were easier the ordinary experts were flexible. In the third experiment, ordinary experts could not work out that the plan they had correctly rejected in the previous problem was now the best solution. Only the super experts resisted the temptation of settling for a well-known but non-optimal solution.

The failure of the ordinary experts to find a better solution when they had already found a good one supports the view that experts can be vulnerable to inflexible thought patterns. But the performance of the super experts shows that 'experts are inflexible' would be the wrong conclusion to draw from this failure. The Einstellung effect is very powerful – the problem solving capability of our ordinary experts was reduced by about three SDs when a well-known solution was apparent to them. But the super experts, at least with the range of difficulty of problems used here, were less susceptible to the effect. Greater expertise led to greater flexibility, not less.

Knowledge, and the expertise inevitably linked to it, were also responsible for both forms of expert flexibility demonstrated in the experiments. The optimal solution was more likely to be noticed immediately, even before the nominally more familiar solution, among some super experts. Hence, expertise helped super experts avoid an Einstellung situation in the first place because they immediately found the optimal solution. Even when experts did not find the optimal solution immediately, expertise and knowledge were positively associated with the probability of finding the optimal solution after the non-optimal solution had been generated first. Finally, when knowledge discrepancy was minimized, as in the third experiment, super experts had sufficient resources to outperform their slightly weaker colleagues. In all three instances, knowledge was inextricably and positively related to expert flexibility.

It is possible to explain why expertise is positively related to flexibility within current theories of expertise and skill acquisition. Although all participants experiencing the Einstellung effect in this study were experts (at least 3SDs above mean), there were skill differences between them. Hence, it is plausible to assume that there were differences in the strengths of memory traces for the optimal solution (Anderson & Lebiere, 1998) or that the acquired knowledge is more differentiated in skilful players as they possess larger and more sophisticated chunks and templates (Gobet & Simon, 1996a). Alternatively, more skilful experts have more frequently experienced the motif featured in the optimal solution and stored more instances of it than less skilful players (Logan, 1988). Similarly, the memory of the less familiar but optimal solution/production in super experts is still strong enough to be accessed in the given time because it has been previously strengthened and/or differentiated through numerous successful applications. Ordinary experts, on the other hand, possess less experience with the optimal solution and the solution is not retrieved or activated in the short time at their disposal. Importantly, the source of expert inflexibility is not motivational. They spent as long looking for a better solution as the super experts. It has more to do with the fact that the memory of super experts is better developed and integrated so that they noticed the optimal solution more easily than ordinary experts.

This contrast between ordinary and super experts may explain the initially surprising result of the study by Lewandowsky and Kirsner (2000). They found that expert bush fire fighters made diametrically opposite predictions about how a fire would spread depending on the context in which the problem was presented. This is like the behaviour of our ordinary experts. Change the context by removing the good move and they had no difficulty finding the better one. In contrast to their ordinary experts, and just like our super experts, Lewandowsky and Kirsner's 'senior expert' was able to solve the contradictory problem that confused all other experts in the study (p. 302), presumably because the context offered sufficient cues for retrieving the critical chunks or templates in LTM.

Although theories of both skill acquisition and expertise can provide explanations for the Einstellung effect, the phenomenon observed in the second experiment where super experts often saw the less obvious (and superior) solution before they saw the more familiar one is more difficult to reconcile with current theories. It is not clear, for example, why one of the instances related with the optimal solution would win the race, except by pure chance, if there are more instances related to the familiar solution (Logan, 1988). Similarly, if the familiar production had been reinforced more often than the optimal, then there is little reason to expect the switch between the familiar and optimal solution displayed by super experts (Anderson & Lebiere, 1998). On the other hand, Template Theory (Gobet & Simon, 1998a) explains this occurrence by the type of pattern recognition it proposes, which is highly sensitive to small details in the context. Thus, our experiments shed light on the issue of the respective merits of mechanisms primarily based on indexing knowledge (such as Chunking Theory and Template Theory), mechanisms primarily based on activation and reinforcement due to success (such as ACT-R), and mechanisms storing previous instances of the problem (instance theory). With super experts, it seems that rare but optimal solutions are so well indexed that they can be retrieved in spite of the presence of typical and well-reinforced solutions.

Chunking Theory and Template Theory can be seen as production systems with a large number of highly specialized productions. This contracts with ACT-R, where productions are more general. The conclusion seems to be that, with experts in a domain such as chess, it is necessary to encode in memory a large number of patterns with fine variations. General productions, even properly tuned by strengthening as in ACT-R, are not enough, as expertise in this domain depends on the ability to pick very fine discriminations between positions. Additional evidence for this idea comes from memory experiments where the board positions were modified by translating parts of the board or by taking their mirror image (Gobet & Simon, 1996b; Saariluoma, 1994). Recall performance was affected by these modification, suggesting that players encode information about the exact location of pieces (not only about their relative location by storing relations of defense and attack), presumably because this enables them to encode the context more precisely in long-term memory. It is likely that ACT-R could capture this result by giving more emphasis to mechanisms storing precise chunks of information.

Instance theory, which also emphasizes memory retrieval in the development of skill, was developed for explaining simple tasks where the interest was more in studying decrease in reaction time than selection among similar but different alternatives. In order to explain the “switch” effect, it seems that the theory should include additional mechanisms enabling similarity-based matching, such as is the case in other exemplar theories. (Chunking/template theories and instance theory are similar in that they both propose a constant storage of the problem situation; however, the former theories propose an incremental storage of these positions, which are coded as chunks, while the latter proposes that the entire problem situation is stored as an instance.)

We have proposed a two-stage model, where a solution is first proposed by memory retrieval; this is then followed by a quick search for a potentially better solution. Our data suggest that super experts are more flexible because they are more efficient in both stages: they are more likely to find the optimal solution right away by memory retrieval; failing so, they are still more likely to find the optimal solution by search. While we have discussed at length mechanisms that could enable (super)experts to retrieve memory information rapidly, we have left somewhat unspecified the means by which "search" is carried out. Previous literature on chess expertise (see Gobet, de Voogt & Retschitzki, 2004, for an overview) has indicated that this search consists of two main components: generation of candidate moves, and look-ahead search of the consequences of these moves. With the type of experts discussed in the current paper, and given the short times involved, the first component is most likely to restrict itself to moves generated by memory retrieval. The second component must be fairly abbreviated, although it is likely that some variations are calculated to some depth, as masters have memorized not only moves but also sequences of moves (Gobet, 1997; Simon & Chase, 1973). In particular with Experiment 1, it is likely that look-ahead search consists mostly in checking that the retrieved solution is indeed valid. With Experiment 3, by contrast, the search behavior is much more important, and memory retrieval gets interleaved with look-ahead processes. As noted above, superior knowledge from super experts makes them more flexible not only because they tend to retrieve the optimal solution first, but also because they can evaluate the consequences of a specific move more precisely and therefore adapt their search behavior. Importantly, according to chunking and template theories, the differences between players of different skill levels, and therefore the interaction we have observed between skill and problem difficulty, are not due to the presence of totally different search strategies, but rather to a graded mixture of decisions made by pattern recognition and slow, heuristic search process (Gobet, 1997).

It is possible that chess may not be the best candidate to demonstrate expert inflexibility. At first sight chess, being an adversarial game, does not belong to the category of consistent mapping tasks where the conditions of the task rarely, if ever, change. Unlike the musician who plays the same sequence of notes each time he plays a piece, with perhaps some variation in interpretation, or a tennis player hitting a serve to the same spot, a chess player, once out of the opening, has to deal with new situations he has never encountered before in that particular form. On the other hand, the rules in chess never change and although new situations arise, recognizing previously encountered patterns, and applying solutions connected to those patterns, is believed to be at the core of chess expertise (Chase & Simon, 1973; de Groot, 1946/1978; Gobet & Simon, 1996a). As Palmeri (1997) put it in his account of automaticity and categorization: "whereas novices seem to rely on slow, conscious, deductive reasoning, experts seem to rely on fast, relatively unconscious processing — the chess master 'sees' the right moves" (p. 346). This automation of processes in the highly skilled can be seen in Grand Masters who play a 'simultaneous exhibition' against 30 or more club level players, beating almost all of them. The Grand Master stops at each board for a second or so, selects a move and moves on to the next one. Very little computation is required. A single look at the board will produce moves that will beat the average club player. As our data suggests, however, players do not have to settle for the first solution that comes to their mind, which they anyway rarely do, and can look further for a better solution. For example, the same Grand Master in the simultaneous exhibition who is usually playing quickly, without much search involved, can stop at any time by a board when he realizes that a detailed examination of the position is necessary.

An important question is whether our findings or implication about experts' flexibility would generalize to other domains/tasks where automatization is more likely to play an important role. There are indications that this is the case. Experienced typists are able to stop typing within a couple of keystrokes of discovering an error although their performance is usually assumed to be automated and therefore not under conscious control (Logan, 1982; Long, 1976; see also Logan & Cowan, 1984). Expert table tennis players are able to adjust the bat during the last 150 ms of ball flight (Bootsma & van Wieringen, 1990). It seems that even in domains that are more susceptible to automatization than chess, experts, at least the best of them, are able to maintain close control of their performance, thus remaining flexible.

Our findings have implications for education and training. Einstellung like phenomena seem to be frequent with experts (Woltz et al., 2000). Singley and Anderson (1989) noticed that they represent the majority of errors in expert text editor behaviour. The serious consequences of the Einstellung effect have been demonstrated by Reason (1990) who showed that many accidents in real life are caused by experienced system users such nuclear power-station controllers being stuck in their well-practiced routines and unable to adopt a new pattern of thought when a system malfunction meant that the familiar pattern was no longer appropriate. Our experts who over-estimated the performance of their colleagues in Experiment 1 and 2 underestimated the power of the Einstellung effect. The dangers connected with the Einstellung-like phenomena are real and present problems even for the most skilled experts.[10]

Nevertheless, fear of a lack of adaptability associated with an increase in skill (e.g., Hesketh, 1997; Lubart, 1994; Sternberg, 1996) is unwarranted. Our ordinary experts were not inflexible when faced with an easy problem (Problem 5 in Experiment 2). They were only inflexible when faced with a difficult problem such as that in Experiment 1 (which was not solved by any of the average players in the 1-solution version). But even here super experts were flexible when faced with this problem. A similar situation was observed in Experiments 2 and 3 where the better players benefited from their larger knowledge base that allowed them to notice optimal ways of dealing with the problems. The way to avoid inflexibility is not less expertise but more.

The training required to produce experts should not be seen as a source of potential problems but as a way to acquire the skill to deal effectively and flexibly with all the situations that can arise in the domain. Creativity is a consequence of expertise rather than expertise being a hindrance to creativity. To produce something novel and useful it is necessary first to master the previous knowledge in the domain. More knowledge empowers creativity rather than hurting it (e.g. Kulkarni & Simon, 1988; Simonton, 1997; Weisberg, 1993, 1999).

In summary, in a series of experiments using super experts, contradictory findings from previous research (Lewandowsky & Kirsner, 2000; Saariluoma, 1990) were clarified and unambiguous evidence for the positive association between expertise and flexibility was provided. A possible theory of expert flexibility in the Einstellung like situations based on expertise and skill acquisition theories was also provided. We showed that ordinary experts can be inflexible, that is unable to resist the temptation of choosing a well-known solution, and that for them 'inflexibility of experts' is reality. However, we also showed that super experts maintain control over their performance, noticing and taking into account even the smallest, and, on first sight, maybe irrelevant, details. Inflexibility of experts is both myth and reality. But the greater the degree of expertise, the more of a myth it becomes.

References

Amabile, T. M. (1989). *Growing up creative: Nurturing a lifetime of creativity*. New York, NY: Crown Publishers, Inc.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89,* 369-406.

Anderson, J. R. (1983). *The architecture of cognition*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Anderson, R. J. (1992). Automaticity and the ACT* theory. *American Journal of Psychology, 105,* 165-180.

Anderson, J. R. & Lebiere, C. (1998). *The atomic components of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Atwood, M. E., & Polson, P. G. (1976). A process model for water jug problems. *Cognitive Psychology, 8,* 191-216.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. New York, NY: Cambridge University Press.

Bootsma, J. R. & Van Wieringen, C. P. (1990). Timing an attacking forehand drive in table tennis. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 21-29.

Bryan, W. L., & Harter, N. (1899). Studies on the telegraphic language. The acquisition of hierarchy of habits. *Psychological Review, 6,* 345-375.

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.

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

Chen, Z. & Mo, L. (2004). Schema Induction in Problem Solving: A Multidimensional Analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 583-600.

Csikszentmihalyi, M., (1996). *Creativity: Flow and the psychology of discovery and invention.* New York, NY: Harper Collins.

deBono, E. (1968). *New Think: The use of lateral thinking in the generation of new ideas*. New York: Harper Collins.

de Groot, A. D. (1978). *Thought and choice in chess. (2nd ed)*. Oxford, England: Mouton. (Originally published 1946)

Delaney, P. F., Ericsson, K. A., & Knowles, M. E. (2004). Immediate and Sustained Effects of Planning in a Problem-Solving Task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1219-1234.

Ericsson, K. A. (1998). The scientific study of expert levels of performance: general implication for optimal learning and creativity. *High Ability Studies, 9*, 75-101.

Ericsson, K. A. (2003). The development of elite performance and deliberate practice: An update from the perspective of the expert-performance approach. In J. Starkes and K. A. Ericsson (Eds.), *Expert performance in sport: Recent advances in research on sport expertise* (pp. 49-81). Champaign, IL: Human Kinetics.

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist, 49,* 725-747.

Ericsson, K. A., & Delaney, P. F. (1999). Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 257-297). New York, NY: Cambridge University Press.

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

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

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (rev. Ed.). Cambridge, MA: The MIT Press.

Feltovich, P. J., Spiro, R. J., & Coulson, R. L. (1997). Issues of expert flexibility in contexts characterized by complexity and change. In P. J. Feltovich, K. M. Ford & R. R. Hoffman (Eds.), *Expertise in context: Human and machine* (pp. 125-146). Cambridge, MA: The MIT Press.

Fitts, P. M. (1964). Perceptual-motor skill learning. In A. W. Melton (Ed.), *Categories of human learning* (pp. 243-285). New York, NY: Academic Press.

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

Frensch, P. A., & Sternberg, R. J. (1989). Expertise and intelligent thinking: When is it worse to know better? In R. J. Sternberg, (Ed), *Advances in the psychology of human intelligence Vol. 5.* (pp. 157-188). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Frensch, P. A., & Sternberg, R. J. (1991). Skill-related differences in game playing. In R. J. Sternberg & P.A. Frensch (Eds.), *Complex problem solving: Principles and mechanisms* (pp. 343-381). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Gobet, F. (1997). A pattern-recognition theory of search in expert problem solving. *Thinking & Reasoning, 3,* 291-313.

Gobet, F. (2000). Some shortcomings of long-term working memory. *British Journal of Psychology, 91,* 551-570.

Gobet F. & Jansen, P. (1994). Towards a chess program based on a model of human memory. In H. J. van den Herik, I. S. Herschberg, & J. W. Uiterwijk (Eds.), Advances in Computer Chess 7 (pp. 35–60). Maastricht: University of Limburg Press.

Gobet, F., Lane, P. C. R., Croker, S., Cheng, P. C-H., Jones, G., Oliver, I. & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences, 5,* 236-243.

Gobet, F. & Simon, H. A. (1996a). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology, 31,* 1-40.

Gobet, F., & Simon, H. A. (1996b). Recall of random and distorted positions. Implications for the theory of expertise. *Memory & Cognition, 24*, 493-503.

Gobet, F. & Simon, H. A. (2000). Five seconds or sixty? Presentation time in expert memory. *Cognitive Science, 24,* 651-682.

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

Gobet, F., & Waters, A. J. (2003). The role of constraints in expert memory. *Journal of Experimental Psychology: Learning, Memory & Cognition, 29,* 1082-1094.

Haider, H. & Frensch, P. A. (1996). The Role of Information Reduction in Skill Acquisition. *Cognitive Psychology, 30,* 304-37.

Haider, H., & Frensch, P. A. (1999a). Eye movement during skill acquisition: More evidence for the information-reduction hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 172-190.

Haider, H., & Frensch, P. A. (1999b). Information Reduction During Skill Acquisition : The Influence of Task Instruction. *Journal of Experimental Psychology: Applied, 5,* 129-151.

Hayes, J. R. (1989). Cognitive processes in creativity. In J. A. Glover, R. R. Ronning, R. R. & C. R Reynolds (Eds.), *Handbook of creativity. Perspectives on individual differences* (pp. 135-145). New York, NY: Plenum Press.

Hecht, H., & Proffitt, D. R. (1995). The price of expertise: Effects of experience on the water-level task. *Psychological Science, 6,* 90-95.

Hesketh, B. (1997). Dilemmas in training for transfer and retention. *Applied Psychology: An International Review, 46,* 317-339.

James, W. (1939). *Talks to teachers on psychology*. Oxford, England: Holt. (Originally published 1899).

Koestler, A. (1989). *The act of creation*. New York, NY: Arkana/Penguin Books.

Krems, J. F. (1995). Cognitive flexibility and complex problem solving. In P. A. Frensch & J. Funke (Eds.). *Complex problem solving: The European perspective*. (pp. 201-218). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Kuhn, S.T. (1996). *The Structure of scientific revolutions*. Chicago: University of Chicago Press. (Originally published 1962).

Kulkarni, D., & Simon, H. A. (1988). The processes of scientific discovery: The strategy of experimentation. *Cognitive Science, 12,* 139-175.

Lewandowsky, S., & Kirsner, K. (2000). Knowledge partitioning: Context-dependent use of expertise. *Memory and Cognition, 28,* 295-305.

Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data using generalized linear models. *Biometrika, 73,* 12-22.

Lippman, L. G. (1996). A short report concerning the Einstellung effect. *Journal of General Psychology, 123,* 233-235.

Logan, G. D. (1982). On the ability to inhibit complex movements: A stop-signal study of typewriting. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 778-792.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492-527.

Logan, G. D. & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91,* 295-327.

Long, J. B. (1976) Visual feedback and skilled keying: Differential effects of masking the printed copy and the keyboard. *Ergonomics, 19,* 93-110.

Lovett, M. C. & Anderson, J. R. (1996). History of success and current context in problem solving: Combined influences on operator selection. *Cognitive Psychology, 31,* 168-217.

Lubart, T. I. (1994). Creativity. In R. J. Sternberg (Ed.), *Thinking and problem solving*. San Diego, CA: Academic Press, Inc.

Luchins, A. S. (1942). Mechanization in problem solving--the effect of Einstellung. *Psychological Monographs, 54, No. 6*, 95.

Luchins, A. S., & Luchins, E. H. (1959). *Rigidity of behavior: A variational approach to the effect of Einstellung*. Oxford, England: University of Oregon Press.

McKelvie, S. J. (1990). Einstellung: Luchins' effect lives on. *Journal of Social Behavior and Personality, 5,* 105-121.

Mednick, S. (1962). The associative basis of the creative process. *Psychological Review, 69,* 220-232.

Minsky, M. (1977). Frame-system theory. In P. N. Johnson-Laird & P. C. Wason (Eds.), *Thinking: Readings in Cognitive Science* (pp. 355-376). New York: McGraw-Hill.

Palmeri, J. T. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 324-354.

Piaget, J., & Inhelder, B. (1956). *The child's conception of space*. London: Routledge and Kegan Paul. (Originally published 1948).

Reason, J. (1990). *Human error*. New York, NY: Cambridge University Press.

Richman, H. B., Gobet, F., Staszewski, J. J., & Simon, H. A. (1996). Perceptual and memory processes in the acquisition of expert performance: The EPAM model. In K. A. Ericsson (Ed.), *The road to excellence* (pp. 167-187). Mahwah, NJ: Erlbaum.

Rickard, C. T. (1997). Bending the power law: A CMPL theory of strategy shifts and the automatization of cognitive skills. *Journal of Experimental Psychology: General, 126,* 288-311.

Rickard, C. T. (2004). Strategy execution in cognitive skill learning: An item-level test of candidate models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 65-82.

Rosenbaum, A. D., Carlson, A. R., & Gilmore, O. R. (2001). Acquisition of intellectual and perceptual-motor skills. *Annual Review of Psychology, 52,* 453-470.

Saariluoma, P. (1990). Apperception and restructuring in chess players' problem solving. In K. J. Gilhooly, M.T. Keane, R.H. Logie & G. Erdos (Eds.), *Lines of thinking: Reflections on the psychology of thought*, *Vol. 2.* (pp. 41-57). Oxford, England: John Wiley & Sons.

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

Saariluoma, P. (1994). Location coding in chess. *The Quarterly Journal of Experimental Psychology, 47A*, 607-630.

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

Saariluoma, P., & Hohlfeld, M. (1994). Apperception in chess players' long-range planning. *European Journal of Cognitive Psychology, 6,* 1–22.

Schneider, W., Dumais, S. T. & Shiffrin, R. M. (1984). Automatic and control processing and attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 1-27). Orlando, FL: Academic Press.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84,* 1-66.

Shiffrin, R. M., & Schneider, W. A. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84,* 127-190.

Simon, H. A. (1969). *The sciences of the artificial.* Cambridge, MA: MIT Press.

Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist, 61,* 393-403.

Simonton, D. K. (1997). Creative productivity: A predictive and explanatory model of career trajectories and landmarks. *Psychological Review*, *104,* 66-89.

Simonton, E. K. (1999). *Origins of genius: Darwinian perspectives on creativity*. Oxford University Press, New York.

Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge, MA, US: Harvard University Press.

Sternberg, R. J. (1996). Costs of expertise. In K. A. Ericsson (Ed), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games* (pp. 347-354). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Sternberg, R. J. (Ed.). (1999). *Handbook of creativity.* New York, US: Cambridge University Press.

Sternberg, R. J., & Frensch, P. A. (1992). On being an expert: A cost-benefit analysis. In R. R. Hoffman (Ed.), *The psychology of expertise: Cognitive research and Empirical AI* (pp. 191-204). New York, NY: Springer-Verlag.

Weisberg, R. W. (1993). *Creativity: Beyond the myth of genius*. San Francisco: Freeman.

Weisberg, R. W. (1999). Creativity and knowledge: A challenge to theories. In. R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 227-250). New York, NY: Cambridge University Press.

Weisberg, R. W. (2006). *Creativity: Understanding innovation in problem solving, science, invention, and the arts.* Hoboken, NJ: John Wiley.

Wiley, J. (1998). Expertise as mental set: The effects of domain knowledge in creative problem solving. *Memory and Cognition, 26,* 716-730.

Wertheimer, M. (1959). *Productive thinking*. Oxford, England: Harper.

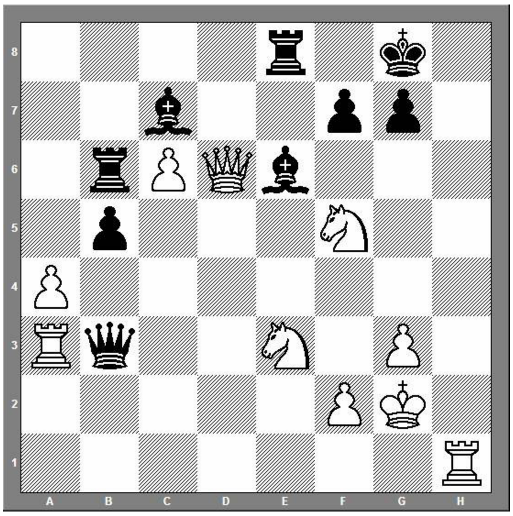
Woltz, D. J., Gardner, M. K., & Bell, B. G. (2000). Negative transfer errors in sequential cognitive skills: strong-but-wrong sequence application. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 601 - 625.

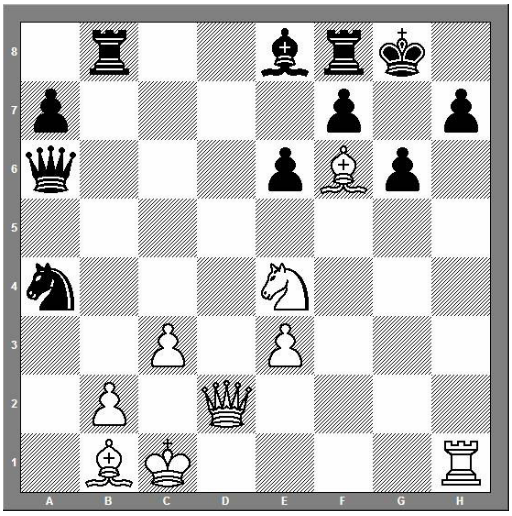
Zeitz, C. M. (1997). Some concrete advantages of abstraction: How experts' representations facilitate reasoning. In P.J. Feltovich, M. K. Ford, & R. R. Hoffman, (Eds.), *Expertise in context: Human and machine* (pp. 43-65). Cambridge, MA: The MIT Press.

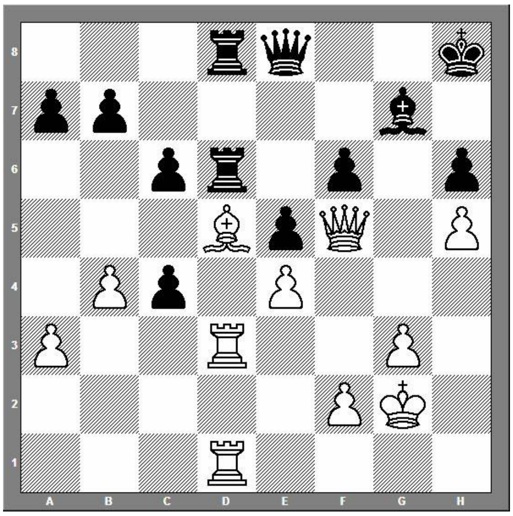
Appendix A

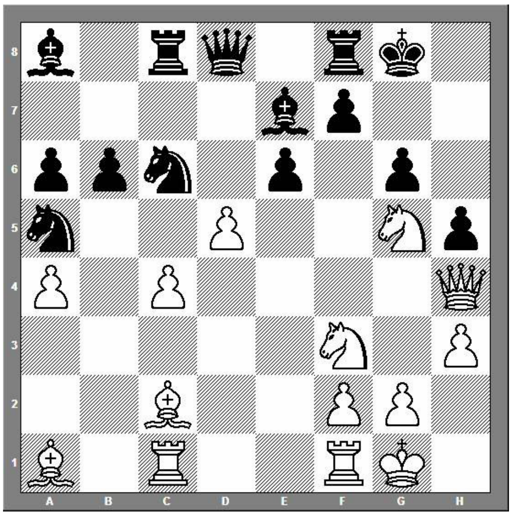
Problems used in Experiment 2. White to move.

2-solution (critical) problems (the first problem can be found in Figure 2):

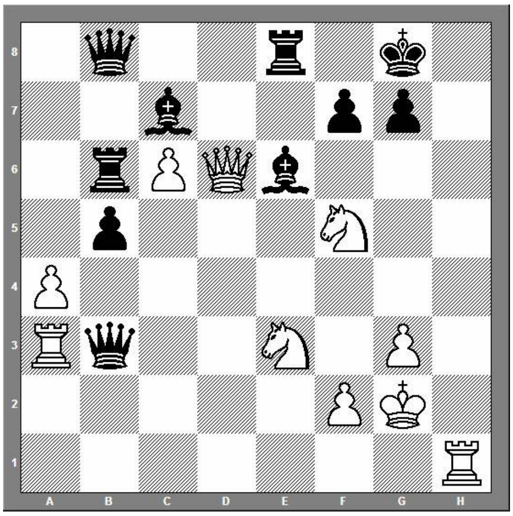
2c) 

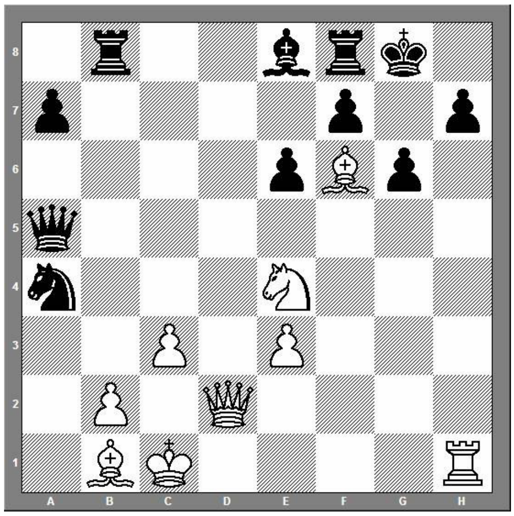
3c) 

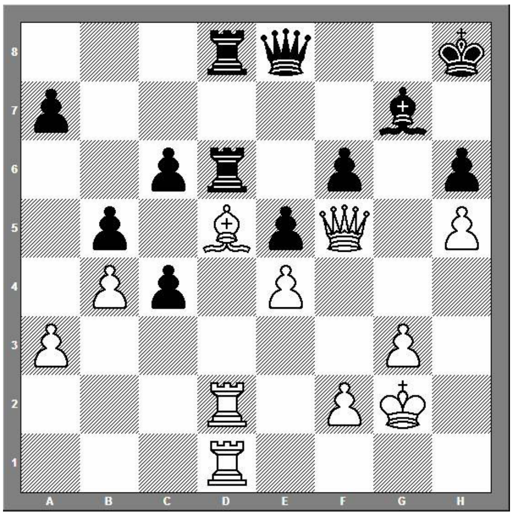
4c) 

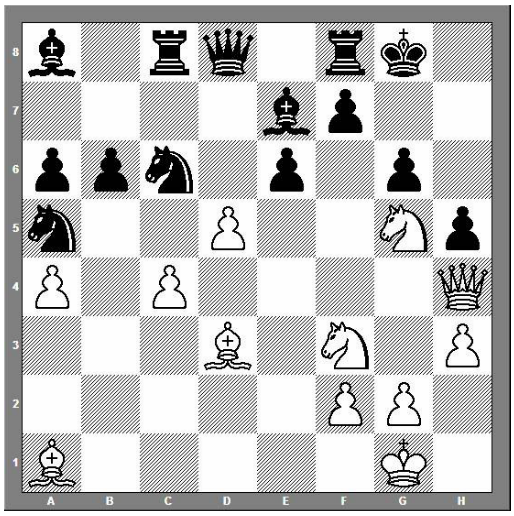
5c) 

1-solution (extinction problems):

2e) 

3e) 

4c) 

5c) 

Solutions: Optimal and familiar (in brackets) solutions: 2c) 1. Qf8 (1. Ne7/Rxb3); 3c) 1. Qh2 (1. Rxh7); 4c) 1. Bg8 (1. Bxc4); 5c) 1. Qxh5 (1. dxc6)

Appendix B

Simplified explanation of the problems used in Experiment 3

Problem 1: In this position Black gets a play on the dark squares with 1…g5. A possible continuation could be: 2. fxg5 Ne5 3. Qe2 Nfg4 followed by capture of the pawn on g5 and good play.

Problem 2: In this position the bishop is on e2 which disables the g5 blow from the first problem: 1…g5 2. fxg5 Ne5 3. Qe3! Nfg4 4. Qd2 Rg8 5. Bxg4! Nxg4 6. Nf3! and White keeps the pawn on g5. Hence, after 3. Qe3 Black has to opt for the passive Nh7 which gives an advantage to White (e.g., 4. Nf3 Nxf3 5. Bxf3 hxg5 6. Bg3 Nf8 7.e5! d5 8. Bxd5! exd5 9. Nxd5 with Nxe7 and Qxg5+). The right move in this problem is 1…. b5 with a complex game.

Problem 3: This time the g5 blow works although the bishop is on e2. The reason is that the knight is now on b3 and cannot come to f3 as in the previous examples to challenge Black's control of e5 square and g5 pawn. Hence, 1…. g5 2. fxg5 Ne5 3. Qe3 and now 3…. Nh7 is acceptable because the knight soon goes to a better place – 4. Kb1 hxg5 5. Bg3 Rc8 6. Rf1 Nf8! 7. Nd2 Nfg6! controlling the dark squares.

Author Notes

Merim Bilalić, Department of Experimental Psychology, Oxford University; Peter McLeod, Department of Experimental Psychology, Oxford University; Fernand Gobet, School of Social Sciences, Brunel University.

This research is based on a doctoral dissertation of the first author who was supported by Oxford University Clarendon and ORS scholarships. The preparation of this paper was enabled by an ESRC Postdoctoral Fellowship to the first author. We would like to thank Francis H. C. Marriot for statistical advice; Richard Palliser for advise and help in choosing and analyzing the problems in Experiment 3; Jill O'Reilly, Ann Dowker, Guillermo Campitelli, Robert Gaschler, Gordon Logan, Eric-Jan Wagenmakers, Pertti Saariluoma, and an anonymous reviewer for helpful comments on earlier drafts of the manuscript. Finally, our sincere gratitude goes to all the chess players who took part in the experiments.

Correspondence concerning this article should be addressed to Fernand Gobet, School of Social Sciences, Brunel University, Uxbridge, Middlesex UB7 3PH, Great Britain. E-mail: fernand.gobet@brunel.ac.uk

Footnotes

[1] It should also be noted that a more controlled replication of the study (Vasta, Rosenberg, Knott, & Gaze, 1997) showed a different result – people with more experience in dealing with liquids were better at solving the water level task.

[2] In this text 'familiar solution' describes solutions that are well-known, more likely to be spotted by the average player than less familiar solutions, and sometimes even obvious.

[3] United States Chess Federation (USCF) rating is comparable to the Elo rating (the official rating of FIDE – the International Chess Association), an interval scale for representing the skill of chess players. The scale has a mean of 1500 and a standard deviation of 200. Players with the rating between 1400 and 1600 are called Class C players, 1600-1800 Class B, 1800-2000 Class A, 2000-2200 Experts, 2200-2400 Masters, 2400-2500 International Masters and above 2500 Grand Masters.

[4] The standard name for players between Elo 2000 and 2200 is 'Experts'. However, in this paper we adopted 'Candidate Masters' in order to avoid confusion with the stronger players in our sample who are undeniably 'experts'.

[5] The instruction in our experiment was to "find the shortest way to win" rather than the more ambiguous "find the best move". It is difficult to argue that the shorter line is 'better' than the smothered mate line. Both give mate, both are forced.

[6] We showed the 2-solution problem to a number of players weaker than Candidate Masters and all of them failed to find the optimal solution. This finding confirms that the success rate in finding the optimal solution is not a U shaped curve.

[7] One could argue that the real rigidity displayed by the participants of Luchins' water jug problem (who failed to solve the 1-solution problem) did not occur with chess experts because they all solved the 1-solution problem (International Masters needed on average 23 s,

Masters 44 s, and Candidate Masters 120 s to find the optimal solution). It would have been, however, most surprising if expert chess players could not find a three-move checkmate.

[8] As in the first experiment, we tested several lesser players (weaker than Candidate Master) on the 2-solution (critical) versions of the first four problems. Although some of the players managed to find the optimal solution in some of the problems, the success rate was far below the level of the Candidate Masters.

[9] The position for the problem was taken from the game Freitag – Danner (Oberwart, 1991), after White's 13th move.

[10] One could argue that the problems used in this study do not show the real dangers of the Einstellung effect as presented by Reason (1990) because the non-optimal (familiar) solutions, although not the best solutions, are nevertheless good enough. While this may be the case with most of the problems in the study, Problem 3 in Experiment 2, where the familiar solution is not merely inferior to the best solution as in the other problems, but leads to a forced loss of the game, is a good example of how dangerous the Einstellung effect in chess can be. (Table 3 shows that the lowest scores, even among super experts, are on this problem).

Table 1.

The classification, mean Elo rating and numbers of participants in each group.

Experiment.	*Classification*	*Mean Elo ± SD*	*SDs above mean*	*n*
1	Grand Master	2554 ± 32	5.5	6
	International Master	2443 ± 24	5	6
	Master	2305 ± 58	4	11
	Candidate Master	2129 ± 45	3	11
	Class A*	1895 ± 63	2	8
	Class B*	1677 ± 67	1	8
	Class C*	1506 ± 62	0	8
2	Grand & International Master	2515 ± 64	5	12
	Master	2301 ± 39	4	12
	Candidate Master	2116 ± 58	3	12
3	Grand Master	2563 ± 50	5.5	5
	Master	2304 ± 56	4	4
	Candidate Master	2146 ± 56	3	3

* These players also constituted the group of weaker players in Experiment 2.

Table 2.

Experiment 1. The percentage of players in Group 1 (experts) who found the optimal solution in the 2-solution (critical) or 1-solution (extinction) problems and the percentage of players in Group 2 (less skilful players) who found the solution in the 1-solution problem. (Grand Masters were not shown the 1-solution problem as they all found the optimal solution in the 2-solution problem.)

	Problem	
Skill Level	2-solution (critical)	1-solution (extinction)
Group 1		
Grand Master	100%	-
International Master	50%	100%
Master	18%	100%
Candidate Master	0%	100%
Group 2		
Class A	-	63%
Class B	-	13%
Class C	-	0%

Note. The players in Group 2 were presented with the 1-solution (extinction) problem only.

Table 3.

Experiment 2. The percentage of expert players (Group 1) who found the optimal solution in the first four 2-solution (critical) problems and the percentage of less skilful players (Group 2) who found the solution to the corresponding 1-solution (extinction) problems.

	Problem				
Skill Level	1	2	3	4	*Mean*
Group 1					
IM and GM	100%	92%	75%	100%	92%
Master	67%	67%	67%	75%	69%
Candidate Master	42%	33%	13%	50%	35%
Group 2					
Class A	100%	88%	100%	75%	91%
Class B	63%	63%	88%	50%	66%
Class C	50%	50%	50%	38%	47%

Note. International Masters (IM) and Grand Masters (GM) were collapsed into a single group because of the small number of players in these two categories.

Table 4.

Experiment 2. The percentage of players who found the optimal solution in the 2-solution (critical) problems immediately.

	Problem				
Skill Level	1	2	3	4	*Mean*
IM and GM	66%	75%	0%	58%	50%
Master	50%	25%	0%	25%	25%
Candidate Master	17%	0%	0%	17%	9%

Note. IM = International Master. GM = Grand Master.

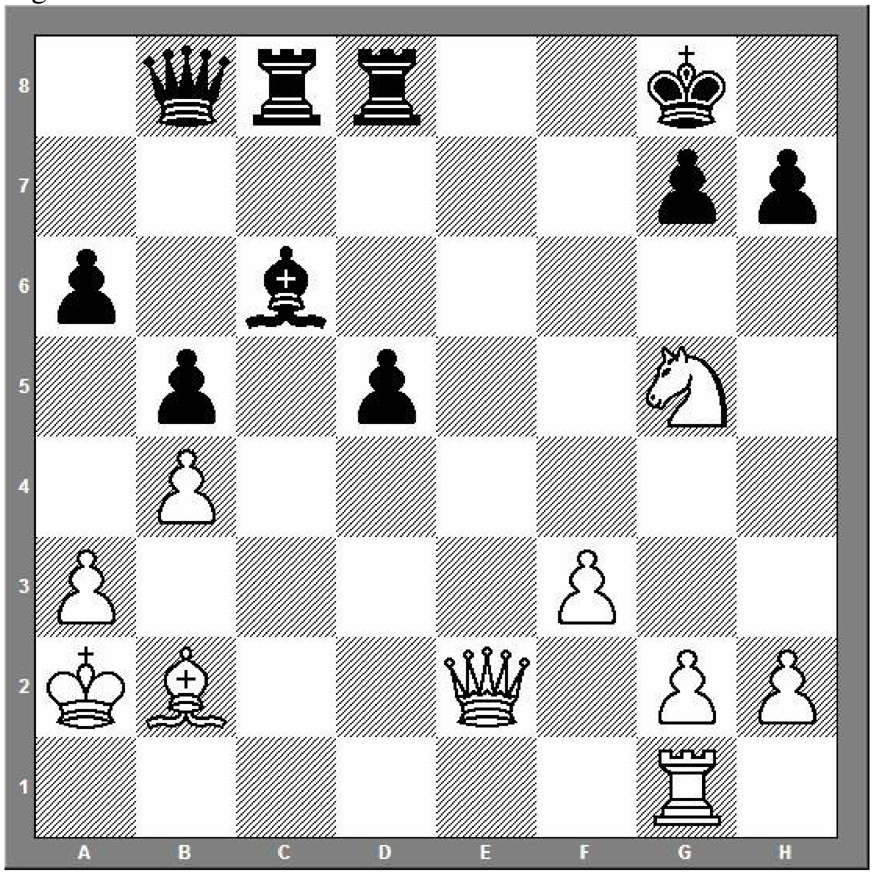
Figure Captions

Figure 1. The 2-solution (critical) and 1-solution (extinction) problem in Experiment 1 (White to play). In the 2-solution problem the well-known smothered mate solution is possible: 1. Qe6+ Kh8 2. Nf7+ Kg8 3. Nh6++ Kh8 4. Qg8+ Rxg8 5. Nf7#. The shorter solution is: 1. Qe6+ Kh8 (if 1…. Kf8 2. Nxh7#) 2. Qh6 Rd7 3. Qxh7#. In the 1-solution problem the black bishop has moved from c6 to h5. This disables the smothered mate solution but still allows the shorter solution. (In the 1-solution problem, playing 2. …Bg6 does not save black because of 3. Qxg7#.)

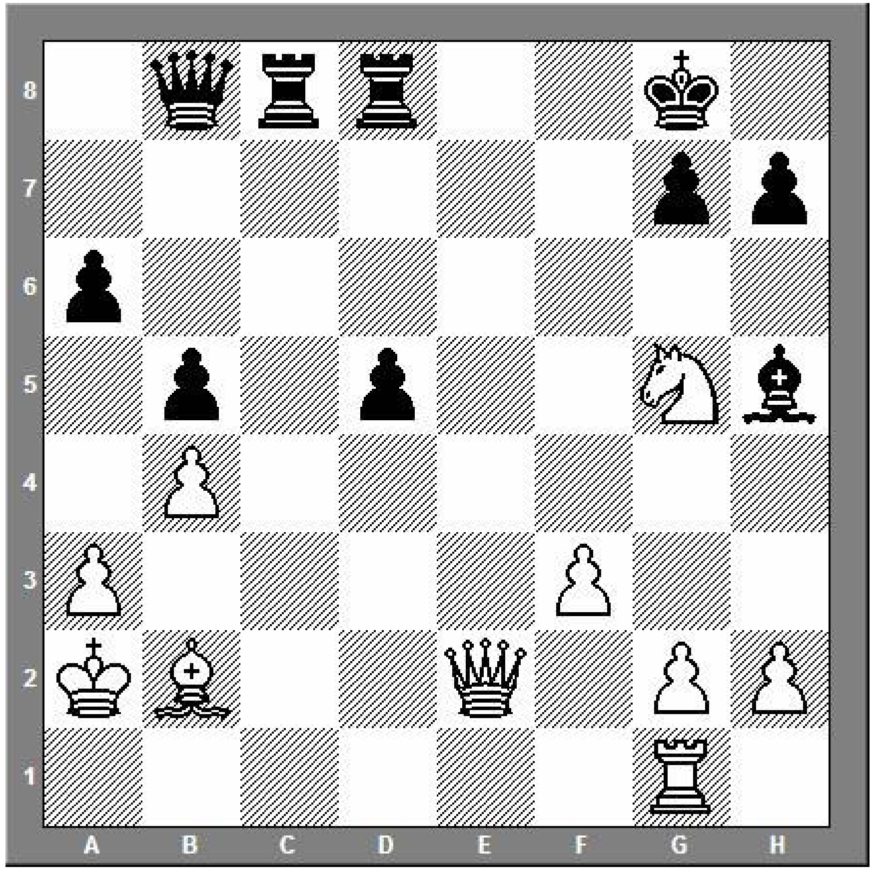
Figure 2. Problem 1 in Experiment 2 (White to play). The familiar solution, 1. Nf8+, which wins at east the exchange (2. Nxd7), is inferior to the less obvious solution – 1. Qxh6+ which forces mate: 1…. Kxh6 2. g5+ Kh7 3. Rh4#.

Figure 3. The three problems in Experiment 3 (Black to play). The correct solution in the first problem is 1…. g5 while 1…. b5 is the best in the second one because the bishop is now on e2 and not on d3 as in the first problem. In third problem 1…. g5 is again the correct solution despite the bishop being on e2. The crucial difference which allows 1…. g5 is that is the knight is on b3 instead of d4 in the previous two problems. A more detailed description can found be in Appendix B while a full analysis can be obtained from the first author.

Figure 1

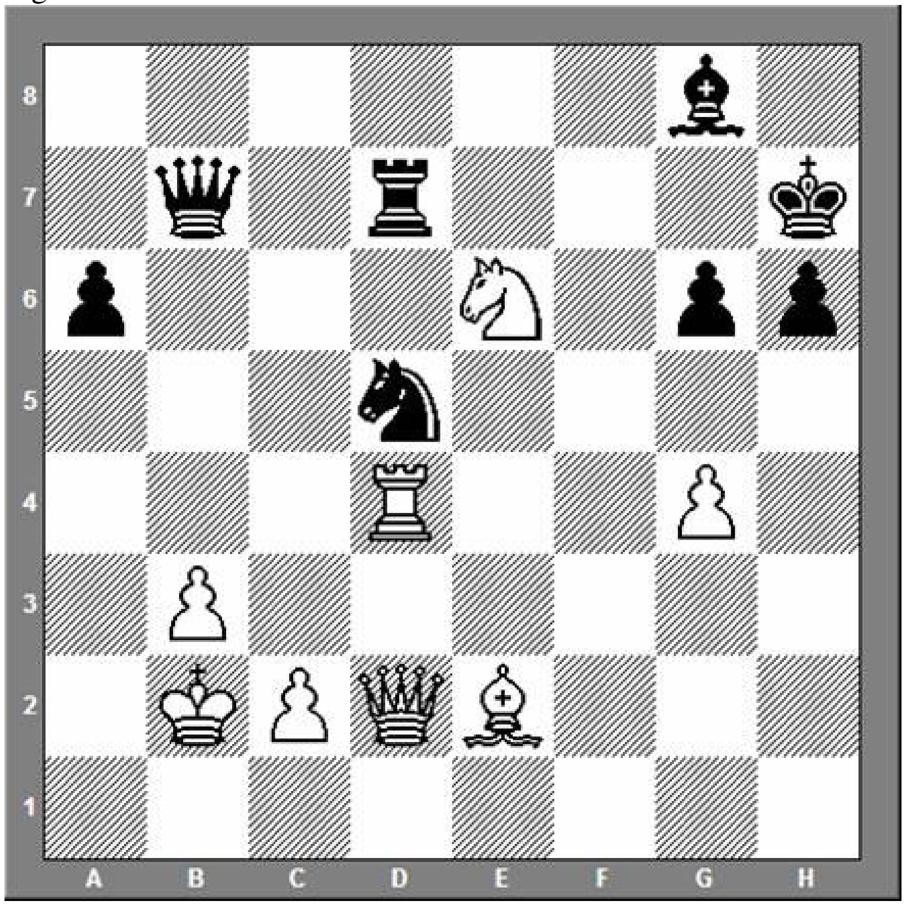


2-solution (critical) problem

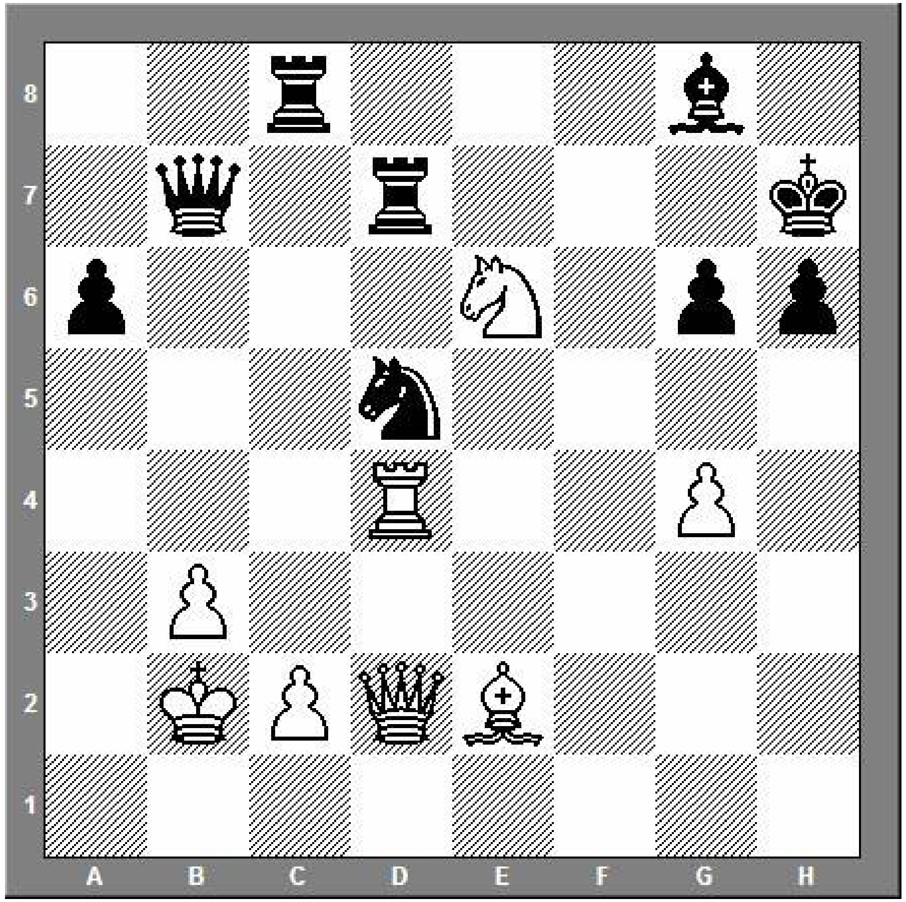


1-solution (extinction) problem

Figure 2

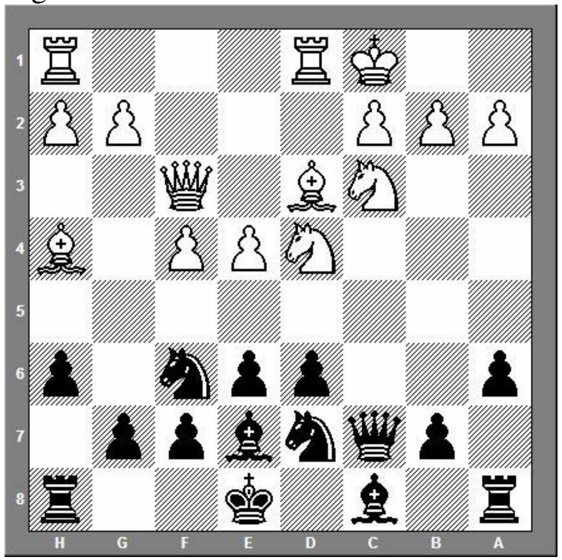


2-solution (critical) problem

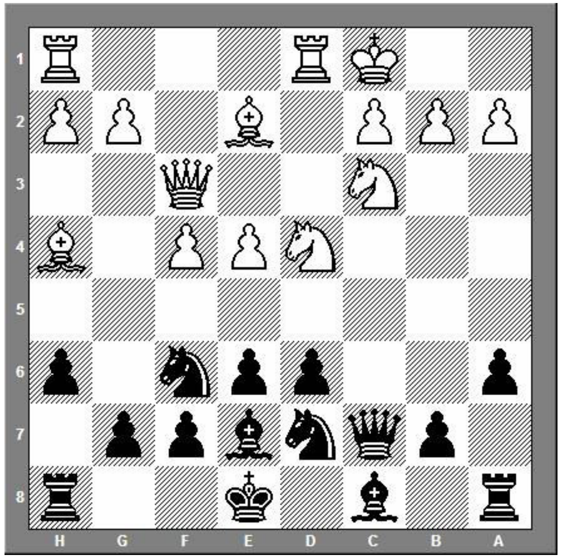


1-solution (extinction) problem

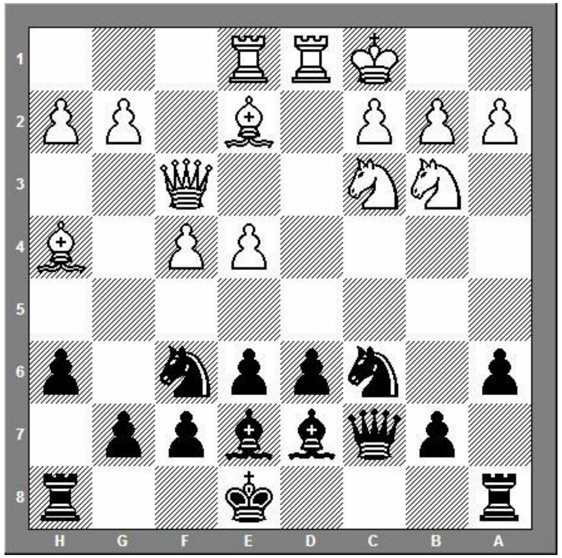
Figure 3



Problem 1



Problem 2



Problem 3