Campitelli, G., Gobet, F., Williams, G., & Parker, A. (2007). Integration of perceptual input and visual imagery in chess players: Evidence from eye movements. *Swiss Journal of Psychology, 66,* 201–213.

Integration of Perceptual Input and Visual Imagery in Chess Players: Evidence from Eye Movements

Guillermo Campitelli
Fernand Gobet
Centre for the Study of Expertise
Brunel University

Gareth Williams
School of Psychology
University of Nottingham

Amanda Parker
Psychology, Brain and Behaviour, School of Biology
University of Newcastle

Send correspondence to:
Fernand Gobet
Centre for the Study of Expertise
Brunel University
Uxbridge, Middlesex, UB8 3PH
United Kingdom
Telephone: +44 (1895) 265484
Fax: +44 (1895) 237573
Email: Fernand.Gobet@brunel.ac.uk

Running Head: Expert visual imagery

Abstract

This multiple case study addresses the question of how information from the environment is integrated with mental images. Chessplayers (N = 4) of different levels were submitted to a visual imagery task, with familiar stimuli (chess positions) and unfamiliar stimuli (boards containing shapes). They were visually presented with a position that remained fixed, and with a grid where moves were displayed using a standard chess notation familiar to the participants. Their task was to mentally reproduce a sequence of moves from the original position. Retention of updated positions was assessed with a memory task. Eye movements were recorded during the entire experiment. We found that (a) players performed better with familiar stimuli than with unfamiliar stimuli; (b) there was a strong correlation between skill level and performance in the familiar, but not unfamiliar condition; (c) players used the external board as an external memory store; but (d) there was no difference in the extent to which players of different skill levels shifted their attention to the external board. Control tasks unrelated to chess established that the skilled players did not differ from the unskilled in general cognitive abilities. These results emphasize the role of long-term memory in expertise and suggest that players use processes that enable them to smoothly combine information from the environment with mental images.

Keywords: blindfold chess, environment, expertise, eye movements, mental images

Integration of Perceptual Input and Visual Imagery in Chess Players: Evidence from Eye Movements

When we perform a mental calculation and someone else is talking nearby, we ask the person to be quieter. When we engage in a “deep” conversation, we switch the television off. In these cases the psychological task we want to carry out is better performed if we disconnect or reduce our sensory connection with the environment. On other occasions, we seem to connect our psychological processes with perceptual input from the environment in order to improve our performance. For example, in order to estimate how some pieces of furniture fit in an empty room, we look at this room rather than imaging it; or, when calculating our monthly budget, we write down figures of our income and expenses on a piece of paper in order to aid our thinking and decision making.

These two phenomena, far from being contradictory, are two sides of the same coin: we are able to modulate our connection with the environment and, hence, how much information we want to receive from it. Sometimes the modulation is a choice and sometimes a necessity, depending on the requirements of the task, its complexity, and the expertise of the individual performing the task. For example, almost all chess grandmasters are capable of playing an entire game (or even several simultaneous games) with their eyes closed, but in a tournament they will usually look at the board while reflecting on the next move to play. By contrast, novice players are not capable of playing a game blindfold and will always look at the board when thinking.

Imagery and Eye Gaze

One way of studying these two phenomena is to record where the participants are looking at while they perform a cognitive task. Glenberg, Schroeder and Robertson (1998) showed that participants averted their gaze from complex stimuli (e.g., the face of the experimenter, a scene, or a film) when they had to perform relatively complex cognitive tasks for which those stimuli were irrelevant. The gaze aversion was a direct function of the level of difficulty of the task.

A more precise measure is offered by eye-movement recording. Hebb (1968) proposed a relationship between eye movements and imagery: if an internal image is a reinstatement of the perceptual processes that were used when seeing the same external image, then these processes must include the eye movements. This idea has received some support only recently. Brandt and Stark (1997) showed that the eye scan-path over a blank 6 x 6 grid during the imagination of a pattern of shadowed squares resembled the eye scan-path during the viewing of the same pattern. Laeng and Teodorescu (2002) asked participants first to look at a stimulus, and second to imagine the same stimulus, while keeping their eyes open and looking at a white screen. During the second phase, the participants had to answer a question about the currently imagined stimulus. Laeng and Teodorescu showed that the performance was better in a condition in which participants were allowed to move their eyes freely both during the viewing and imagery phases than in a condition in which the participants moved their eyes freely only in the viewing phase. This result suggests that the eye movements over a blank screen during imaging are functional for remembering the details of a stimulus.

Spivey and Geng (2001) aurally presented sentences describing horizontal or vertical movements (e.g., a person walking in a train, or events happening in different floors of the same building) while the participants were looking at a blank screen. They showed that the participants performed more horizontal eye movements during the “horizontal” sentence than during the “vertical” and vice-versa. In a second experiment the participants viewed a set of four objects displayed on a white screen;

then, one of the objects disappeared and the participants had to answer a question regarding the missing item. Spivey and Geng (2001) showed that the participants were prone to fixate on the part of the screen where the missing object was before. Moreover, participants were more likely to fixate the blank region when the spatial context was made richer during the test phase by presenting a frame on the screen. The effect was even stronger with the presentation of a grid.

Chess Expertise

Another way of studying those phenomena is offered by chess. In a chess game, when players are thinking which move to play in a given position (e.g., whether moving the king is a better move than moving a pawn), they have to generate sequences of moves internally, and evaluate the position obtained at the end of such sequences. During all this time, the position remains unchanged on the external board. The difficulty of the task is compounded by the fact that the search will sometimes fork in two or more branches. Thus, finding a good move will require a search through the space of possible moves, with frequent backtracking to previous positions—with the obvious requirement that some memory of the positions visited during search must be held, presumably as mental images. In part because of these features, chess has attracted a large amount of research in psychology (Charness, 1992; Gobet, De Voogt, & Retschitzki, 2004; Saariluoma, 1995).

There is a substantial literature showing that expertise in chess and in other domains is acquired by storing domain-specific patterns in long-term memory, as proposed by pattern-recognition theories of expertise (Chase & Simon, 1973a, Gobet & Simon, 1996a). A classical result is that there are strong skill effects in memory tasks with meaningful domain-specific material, but that these effects nearly disappear when the material becomes less familiar (e.g., through randomization;

Charness, 1976; Chase & Simon, 1973a; Gobet & Simon, 1996b). In line with these results, research on problem solving (Campitelli & Gobet, 2004; Charness, 1981; Gobet, 1998) has shown that measures of search behaviour—such as depth of search, number of positions visited, and speed of search—vary as a function of skill. There is also evidence suggesting that search involves the use of imagery (Gobet, 1997, 1998; Holding, 1985).

Chess Imagery and Blindfold Chess

Informal observations of chess grandmasters show that some of them avert their gaze from the board while thinking their next move in certain situations (most notably grandmaster Vassily Ivanchuk). In fact, strong players are able to play one or several games "blindfold," without seeing the board and the pieces.[1] This makes chess a particular interesting domain for studying mental imagery, both theoretically and empirically. Theoretically, mental imagery has often been proposed as a link between pattern recognition and move selection, for example in Chase and Simon's (1973a, 1973b) chunking theory (hence the title "The mind's eye in chess" in one of their papers). More recently, the mechanisms of pattern recognition, forward search and mental imagery have also been integrated in a modification of the chunking theory, the template theory (Gobet, 1997; Gobet & Simon, 1996a).

Empirically, Milojkovic (1982) and Gruber (1991) carried out experiments investigating the time needed to mentally generate moves. They found that the time to move a piece in the mind's eye negatively correlated with chess skill. Both studies found that how far a piece moved increased reaction time for the novices, but the results were inconsistent with the Masters. Similarly, there were inconsistencies with whether moving a piece diagonally takes more time than moving a piece horizontally or vertically (Charness, 1991; Church & Church, 1977; Milojkovic, 1982).

Alfred Binet (1894) was the first to study blindfold chess scientifically. He asked reputed chess masters to fill in a questionnaire enquiring about the type of representations they used while playing blindfold chess. Contrary to his expectations, the results indicated that skilled players do not encode the perceptual characteristics of the pieces and board, such as the color or style of pieces, but that they used a more abstract type of representation. A modified version of blindfold chess (for scientific purposes) is to present a game to players and ask them to perform memory tasks that require the generation of an internal representation of the game. The moves are either dictated aurally or—slightly stretching the meaning of "blindfold"—displayed visually on a screen; in this case, only the last move, and not the entire board, is visible.

Research has shown that blindfold chess involves a combination of memory, search behaviour, and mental imagery (Bachmann & Oit, 1992; Binet, 1894; Ericsson & Staszewski, 1989; Milojkovic, 1982), where the ability to update mental representations of positions plays an essential role. Saariluoma (1991) and Saariluoma and Kalakoski (1997, 1998) took advantage of these unique features of blindfold chess to carry out a series of important experiments, where the moves were presented either aurally or visually. It was found that blindfold chess relies mainly on visuo-spatial working memory, and makes little use of verbal working memory. Another result was that differences in long-term memory knowledge, rather than differences in imagery ability *per se*, are responsible for the observed skill differences. With dictated games, masters showed an almost perfect memory of actual games, or when the moves were random, but legal. However, their performance dropped drastically when the games consisted of (possibly) illegal moves.

Campitelli and Gobet (2005) used a blindfold presentation similar to that of Saariluoma and Kalakoski (1997), in that only the moving piece was presented visually, but not the entire position. Irrelevant information was always in view, and consisted of the board containing a position unrelated to the game that players were following. They found that chessmasters were excellent at filtering irrelevant perceptual information while they were updating and maintaining a visual image of a game. Only in the case where the irrelevant information kept changing and was not familiar were chessmasters unable to filter it. The results suggest that players used the internal representation of the external board as an aid to update the game they were following.

Preview of Study

In the present study, we combine the two lines of research on imagery we have just reviewed: eye-movement recording in memory tasks and blindfold chess. The use of chess enables us to study the role of domain-specific knowledge in mental imagery, and in particular how this knowledge interacts with the familiarity of the material. The main experiment of this article required chessplayers of different levels to perform a memory task that engaged the use of imagery, the main theoretical question being whether players utilized the external board as an "external memory store," to employ Richardson and Spivey's (2000) term. We used a complex task both with familiar chess stimuli and unfamiliar stimuli where chess pieces were replaced by shapes. Players were visually presented with a position (containing either chess pieces or shapes) that remained fixed and a grid where moves were displayed. The task was to mentally keep track of the sequence of moves starting from the original position. The retention of updated positions was assessed with a memory task, where players had to indicate which piece, if any, was located on a specific square indicated in a "test box."

Thus, three key elements were part of the visual display (the board, the grid, and the test box), and eye movements were recorded during the entire experiment. In order to establish that our participants were representative of players of their skill level, we also gave control tasks that were either related or unrelated to chess. Based on the extant literature, we made three predictions.

As we have seen above, pattern-recognition theories of expertise propose that perceptual knowledge stored in long-term memory mediates expert behaviour. Thus, they clearly predict that there should be a strong correlation between skill and performance in the chess condition but not in the shape condition (Hypothesis 1), because patterns can be recognized in the former condition but not in the latter. A corollary of this hypothesis is that the skill effects discussed in hypotheses 2 and 3 below should occur only in the chess condition, but not in the shape condition.

We have discussed above Campitelli and Gobet's (2005) experiment that led to the proposal that players integrate perceptual input with mental images by using the board as an external memory store. However, because relevant and irrelevant information was presented at the same place within the visual display (i.e., on the board) in that study, one might argue that players did not have a choice but to look at it. In the present experiment, the critical information is presented visually in two different places (grid and text box; see Method below), which do not coincide with the board. The board itself contains a position that remains unchanged, in spite of the game being played, and thus increasingly becomes irrelevant. If players of all levels make use of an external memory store, as suggested by Campitelli and Gobet (2005), they all should repeatedly look at the board with the unchanged position (Hypothesis 2).

Assuming that the previous hypothesis is supported, an important question is whether the frequency of glances at the board is affected by the skill level or the familiarity of the stimuli. As reviewed in the introduction, stronger players can search deeper and more rapidly. Taken together with the fact that strong masters can also play blindfold, a reasonable conjecture is that they do not need to look at the external board as often as weaker players. This leads to the prediction that there should be a strong negative correlation in the chess condition between skill level and the number of glances at the board (Hypothesis 3). As a corollary, in the chess condition, there should also be a strong negative correlation between the number of glances and performance.

In this experiment, we controlled for individual differences by submitting the participants to a large number of tasks, both within and outside their domain of expertise. The logic was that we could estimate the representativeness of our sample by comparing its performance, skill level by skill level, to multiple experiments previously published in the literature. Thus, while the standard strategy in psychology is to assume that individual differences will cancel out in a large sample, so that the group average is representative of the population at large, our strategy is to directly verify that our participants fit the pattern of results considered typical in the literature. This approach is not new, and is common in neuropsychology and psychophysics. It seems particularly appropriate for research into expertise, where large within-domain differences are often found between individuals of various skill levels—as was the case in this experiment. It was used for example in the classic works of Chase and Simon (1973a, 1973b) and Chase and Ericsson (1982).

Experiment 1

Method

Participants

Four male chessplayers participated in this experiment: one international grandmaster (GM) with 2550 Elo points[2] (age: 21 years), one international master (IM, 2500, age 22), one candidate master (CM, 2100, age 19), and one class B player (CB, 1750, age 19). Although relatively small, the difference between GM and IM can be considered as reliable as it has remained stable since the data were collected: in the 17 Elo lists that were published since, the average Elo was 2580 for GM (*SD* = 22.5) and 2526 for IM (*SD* = 22.2).

Material

Eye tracker.

The eye tracker used was an ISCAN RK-726PCI Pupil/Corneal Reflection Tracking System. It consisted of a video-based, dark-pupil-to-corneal-reflection method to track eye movements. It sampled at a rate of 60Hz, and was able to track a subject's eye position with accuracy typically better than 0.3° over a +/- 20° horizontal and vertical range. The camera and infrared light of the eye tracking device were situated midway between the CRT VDU and the subjects' eye (420 mm from the eye) but sufficiently low down so that the view of the screen was unobstructed (approximately 30° from the line of sight perpendicular to the screen). Head movements were restricted with a chin rest and a head restrainer. At the beginning of each experiment, the experimenter asked the participants to direct their sight alternately to the four extremes of a template, as well as to its centre. In this way, the location of the template in term of screen coordinates was known, and the space of the participants' fixations was determined by these borders.

Visual display.

Three elements were part of the visual display (see Figure 1 for the visual display of the chess condition and Figure 2 for the visual display of the shape condition): (a) a board displaying a chess or a shape position (14° x 14° of visual angle), (b) a grid of 16 columns x 9 rows, at the bottom of the screen (41° x 11°), and (c) a blue test box on the left-hand side (6° x 6°). During each game, the position on the board remained fixed; only the information presented on the grid and on the test box varied. The numbers on the columns of the grid indicated the move number; the numbers of the rows identified a new sequence of moves (see below). Note that the same numbers appeared on two rows. This is because the moves for white and black were differentiated. White moves were presented on the rows with the number in red (grey in the Figure) on a white background, and black moves were presented on the rows with a white number on a black background.

INSERT FIGURE 1 ABOUT HERE

INSERT FIGURE 2 ABOUT HERE

Procedure

The two conditions (chess and shape) were blocked. For each condition, three games were presented sequentially. The three chess positions were presented first and the three shape conditions were presented afterwards. The total time of the experiment was 30 minutes. Within each game, there were three phases: initial inspection, move presentation, and test. The initial inspection was clearly differentiated from the other phases. The move presentation and test phases occurred in sequence (see Figure 1). Before discussing the differences between the chess and shape conditions, we explain the chess condition in some detail.

Initial inspection phase.

Each game started with the 5-s presentation of a legend with the game number (1, 2, or 3) and condition type (chess or shape) at the centre of the screen (e.g., "game 2 - chess"), after which a fixation cross was presented for 5 s. Then the visual display, containing a position on the board, appeared. (This position remained unchanged and in view for the entire game.) Participants were allowed to inspect the position for 20 s.

Move presentation phase.

After the initial period of 20 s, a sequence of moves was presented on the grid in algebraic notation,[3] with a standard symbol denoting the type of piece moving. Each move was shown for 3 s; when a new move was presented, the previous move disappeared. White moves were presented on the rows with a red (grey in the Figure 1) number in a white background and black moves on the rows with a white number in a black background. Since the position on the board remained static for the whole game, players had to generate an internal image of the updated positions by following the moves presented on the grid.

During the move presentation phase, the sequence of moves began to go forward from move 1 until a test phase started (see "test phase" below and Figure 1 for an illustration of the ordering of the move presentation and test phases; see also the Appendix for the first chess game, showing how the sequence of moves is intertwined with the sequence of test phases). In some games, an arrow indicated a shift to a previous position. For example, Figure 1 shows three oblique backward arrows indicating to the players that they had to recover the representation of the position at a previous stage of the game. After such a shift, one of two things happened: (a) a new test phase started, or (b) a new move presentation phase started. When (b) happened, a new sequence of moves—going forward from that position—

was presented on the two rows underneath the ones that were used for the previous sequence. After a number of moves were presented, another test phase began. The numbers of moves presented in the chess condition were 31, 43, and 44 for game 1, 2 and 3, respectively. The same numbers apply to the shape condition.

Test phase.

At some point in the sequence of moves, the grid became blue, indicating that a test phase had commenced (see Figure 2). Using the algebraic notation, the names of six squares (e.g., “c7”) were presented sequentially, each for 3 s, on the test box. The players had to say out loud the name and colour of the piece that was currently on the square indicated in the test box, or say “nothing” if they thought there was no piece. The experimenter wrote down the replies. For example, if the test box showed c7, and at that stage of the game there was a white bishop in the square c7, the correct answer would have been “white bishop”. Once the test phase had finished, a new sequence of moves was presented on the grid, starting from the latest imagined position. After the last test phase, the game finished and a new game started, following the cycle explained above. There were four test phases in game 1, six in game 2, and six in game 3, both in the game and in the shape conditions. Given that there were 6 tests in each test phase, there were 24 (4 test phases times 6 tests), 36, and 36 tests in each of the games, respectively.

Shape condition.

The procedure of the shape condition was exactly the same as that of the chess condition. In particular, the shape games had the same overall structure as the chess games (see Figure 2 for a board displaying a shape position). There were some differences aimed at affecting mental images, which we describe now. The initial shape positions were obtained by randomly replacing the pieces of a chess position with geometrical shapes (squares, triangles, circles, rhombuses, vertical rectangles, and horizontal rectangles), while keeping the original colour of the pieces. During the move presentation phase, the moves on the grid were presented using the corresponding geometrical shapes. Finally, the moves were random (i.e., they did not follow the type of movement of any chess piece) in the shape condition, with the constraint that the average length of moves (measured as number of squares) matched the average length of moves in the chess condition.

Statistical analyses

We performed correlational analyses across the four subjects for each of the games. We were interested more in the consistency between the three games than in the significance level of the individual correlation coefficients. For example, if a statistically non-significant correlation coefficient of .85 occurred in each of the three games, we considered the correlation as significant, because it was replicated twice. On the other hand, if a statistically significant correlation coefficient of .91 appeared in one game while the coefficients in the other games were very different (e.g., .12 and -.34), we did not consider the correlation sufficiently robust.

We used this type of statistical analysis for all the variables investigated in the main experiment of this study. Similarly, in order to be consistent across tasks, we carried out the same analysis with the additional control tasks (see “additional data”). In all cases, the significance level was .05. We used one-tailed tests since we had clear hypotheses about the direction of the effects. In order to compare the performance between the chess and shape conditions, we performed a paired-sample t test.

Results

Table 1 displays the players’ performance in each game for both conditions. The overall performance was 53% (*SE* = 13.1%) in the chess condition and 23.7%

(*SE* = 3.0%) in the shape condition (*t*(3) = 2.75; *p* <.05, one-tailed). With respect to skill effects, Table 1 shows that the correlation between chess skill (Elo) and performance was significant in all three games of the chess condition, but in none of the games in the shape condition. The lack of significance in the shape condition cannot be attributed to a lack of power, because the *r*s were not consistent across the games (negative, almost zero, and positive, respectively). This result supports our first hypothesis—that skill and performance should correlate strongly in the chess condition but not in the shape condition. Incidentally, Table 1 shows an increase in performance in all players from game 1 to 3 in both conditions.

INSERT TABLE 1 ABOUT HERE

The second hypothesis states that players look at the board to use it as an external memory store, both during the move presentation phase, where information is presented in the grid, and the test phase, where information is presented in the test box. We calculated the number of *shifts* between the grid and the board (vertical shifts) and between the test box and the board (horizontal shifts). We defined a shift as fixating on another element after fixating on one of the three elements (grid, board, and test box). The fixations that occurred on intermediate empty areas of the visual display were excluded from this analysis.

If players did not use the board (or used it only occasionally) as an aid to follow the game, then no vertical shifts would appear, or perhaps a few, by chance. By contrast, if they constantly used the board to follow the game, numerous vertical shifts would be observed. Similarly, during the test phase, the number of horizontal shifts would be low in the former case and high in the latter case. Since the absolute number of shifts is irrelevant for this discussion, in each game we calculated the ratio of the number of vertical shifts to the number of moves multiplied by two, and the ratio of the number of horizontal shifts to the number of tests multiplied by two. (If there is a shift from the grid (or test box) to the board, there is necessarily a shift back from the board to the grid (or test box) to carry on doing the task. That is why the number of moves or tests was multiplied by two.) Using these calculations, our second hypothesis was that, for all players and in both conditions, both ratios should be equal to or greater than 1—that is, there should be on average one shift per move or test, or more.

Table 2 shows these ratios for the vertical and horizontal shifts. With the exception of CB with vertical shifts in the shape condition, all players carried out more than one shift per move or test on average. This result supports our second hypothesis. This table also shows that, in the chess condition, there were no strong negative correlations between vertical shifts and skill, vertical shifts and performance, horizontal shifts and skill, or horizontal shifts and performance. This counts against our third hypothesis. In the shape condition, we did not expect any correlation. However, surprisingly, the vertical shifts correlated with skill.

INSERT TABLE 2 ABOUT HERE

Additional Data

It may be argued that IM and GM have stronger visuospatial and imagery skills (rather than larger domain-specific knowledge) than the other players, which would explain the correlations between skill and the other variables. While the standard approach in psychology is to control for individual differences by using large samples, we followed an alternative methodology, as noted at the end of the introduction. Specifically, we controlled for individual differences by testing participants in a large number of tasks related or unrelated to chess.

To be consistent with previous studies, the chess tasks should show that the four players are representative of their skill level, and the non-chess tasks should show that there are no skill differences in tasks unrelated to the domain of expertise. For each participant, we collected additional data on four chess-related measures (quick problem solving, long-term memory for positions, short-term memory for positions, and practice) and four non-chess-related tasks (simple reaction time, long-term memory for photographs, memory span, and visual short-term memory).

Method

Chess Data

Quick problem-solving.

Players had to find the best moves in briefly presented positions. Forty-nine positions of medium complexity were selected from Livshitz (1988). Each position was presented on a computer screen for 5 s, followed by a 5 s black screen. Participants had to verbalize their move within this 10-s period. Given the short times involved, this task taps pattern recognition more than search abilities,

Long-term memory for game positions.

In this task, which was also computer administered, the participants first completed a learning phase in which they were presented with a list of 250 chess positions for 5 s each. Then, they went through four recognition sessions (immediately after the learning phase, six hours after, 24 hours after, and six days after). In each test session 100 stimuli (50 old and 50 new) were presented, and for each one the participants had to decide whether the stimulus presented was part of the original list of 250 positions or not.

Short-term memory for game positions.

We used De Groot's (1978) classic recall task. We presented three types of positions: standard, quiet and complex. Quiet positions are more likely to contain familiar structures than standard positions, and the same applies to standard positions compared to complex positions. Therefore, performance was expected to follow the order quiet, standard, and complex condition. There were four positions for each type. Each position was shown on a computer screen for 5 s, and there was no time limit for reconstruction, which was also done on a computer (see Gobet & Simon, 1998, for details of the software used). Note that the long-term and short-term memory tasks for chess positions directly test the validity of chunk-based theories.

Amount of practice.

For estimating the amount of practice, we asked our participants to fill in a grid asking them how many hours per week they had spent either studying or practicing chess in each year of their chess career (for a similar approach see Charness, Krampe & Mayr, 1996; Charness, Tuffiash, Krampe, Reingold & Vasyukova, 2005; Gobet & Campitelli, 2007). With this information, we calculated the cumulative time spent studying and playing chess.

Non-Chess Data

Simple reaction time.

This task was aimed to control for inter-individual differences in domain-general perceptual and motor abilities. A red or a yellow circle was presented on the screen, and participants had to press the relevant button as fast and accurately as possible. We carried out four sessions of 50 trials each.

Long-term memory for photographs.

This task was the same as long-term memory for chess positions, but using everyday photographs.

Memory span.

We performed three memory span tasks (letter, digits and spatial) with two versions in each (simultaneous and sequential). The letter and digit simultaneous spans started with presenting the participants with a list of 2 items on a computer screen for 2 s. When the items disappeared the participants had to say out loud the series of items in the correct order. There were two trials and if the participants were correct in both, a new series with an additional item was presented and the presentation time increased one second per item added. The memory span was defined as the maximum number of items that can be reproduced correctly in two trials. In the simultaneous spatial task, a number of red squares were presented in a 5 x 5 grid and the participants had to reproduce the location of the squares in a given grid drawn on a piece of paper. In the three cases, the sequential version of the task consisted in presenting a series of items in sequence for 1 s each. The score in each memory span was the average of the simultaneous and the sequential versions.

Short-term memory task for non-game positions.

This task was almost the same as the short-term memory task for chess positions, with the following differences. There were three types of positions: random positions (i.e., a chess position where the location of the pieces on the board had been scrambled), shape-random positions (the same as random chess position, with the addition that the chess pieces were replaced by geometrical shapes) and shape-game (a chess position where the pieces were replaced by geometrical shapes, as in the main experiment). We used 4 positions for each type.

Results

Chess Data

As shown in Table 3, there were significant correlations between skill level and all chess-related measures. These results are in line with previous studies: quick problem solving (e.g., Gobet & Retschitzki, 1991; Charness, 1981); long-term memory (e.g., Gobet & Simon, 2000); short-term memory (e.g., Chase & Simon, 1973a; De Groot & Gobet; 1996; Gobet & Clarkson, 2004); and practice (e.g., Charness et al., 2005). In the memory tasks, performance followed the ordering quiet, standard, and complex, as predicted.

INSERT TABLE 3 ABOUT HERE

Non-Chess Data

Consistent with the literature (Gobet & Simon, 1996a, 2000; Gobet & Clarkson, 2004; Waters, Gobet & Leyden, 2002; see also Masunaga & Horn, 2000, for similar results with the game of Go), Table 4 shows that none of the correlations of the non-chess-related tasks were significant. Two of the three correlations larger (in absolute terms) than .80 either suggest that *weaker* players reacted faster in the simple reaction time task or had a larger spatial memory span. The third correlation does favour the stronger players in the short-term memory for random positions, a task that is not totally unrelated to chess.

In summary, excluding the correlations related to practice, we found that eight out of the eight correlations related to the chess tasks were greater than .90. By contrast, none of the 14 correlations related to the non-chess tasks were greater (in absolute terms) than .90. This clear pattern of results—all positive correlations larger than .90 in the chess-related tasks and none in the non-chess-related tasks—supports the hypothesis that our small sample was a good representation of chessplayers of similar skill levels.

INSERT TABLE 4 ABOUT HERE

Discussion

In this study, we were interested in how chessplayers combine external information with mental images. In particular, we wished to test three specific hypotheses. Our data support the first hypothesis—that there should be a correlation between skill and performance in the chess condition but not in the shape condition. Together with the additional data, where there were strong correlations between skill and performance in chess-related tasks only, this result is in accord with previous studies showing that experts' better performance is limited to domain-specific material (see Gobet et al., 2004). Thus, this pattern of results supports chunk-based theories of expertise, which emphasize the role of long-term memory in memory (Chase & Simon, 1973a; Gobet & Simon, 1996a), problem solving (Campitelli & Gobet, 2004), and imagery (Campitelli & Gobet, 2005) tasks. That is, the existence of a long-term domain-specific knowledge base is the main factor explaining differences in domain-specific tasks.

Our second hypothesis, based on Campitelli and Gobet's (2005) results in an imagery experiment that did not use eye-movement recording, was that all players would follow the strategy of using the external board as an external memory store, regardless of skill differences. This prediction was supported by the data. This result is also in agreement with Spivey and Geng's (2001) and Richardson and Spivey's (2000) memory and imagery experiments, where participants presented with a visual display missing an item looked at the position where this item was located earlier. Given that their effect was made stronger by adding a frame, and in particular a grid, to the display, it is not surprising that we found the same behaviour with our weaker players in a task employing a chessboard. However, it is surprising that the strong players followed this strategy, in spite of the fact that they were able to play blindfold chess and could calculate deep variations in problem solving tasks. The question of skill differences in behaviour was taken up in the third hypothesis.

Laeng and Teodorescu (2002) suggested that, during fixations, participants attempt to reconstruct a mental image. In our experiment, players had to generate an image of a piece moving to a new location and weaken its image in the previous location. In order to facilitate this task, looking at the board seems a sensible thing to do: the new location of pieces can be refreshed, and the perception of the board may help to generate the frame of the internal image with the updated position. Given that stronger players are more capable of generating an internal image of the position without the aid of the external board, we predicted (hypothesis 3) that they would use it less than the weaker players, that is, that the number of shifts would be negatively correlated with skill. Strong players were expected to occasionally use the board as an external aid in case of difficulties (Glenberg et al., 1998), but not as a standard strategy. This prediction was not supported by the data. One possible explanation is that while the strong players could have done the task the way we expected, they chose to do it otherwise, perhaps because looking at the board diminished their cognitive strain. This could be tested in a future experiment in which our task is compared with a new task where the moves are presented on a grid without showing the board simultaneously. Another explanation relates to Fine's (1965) suggestion that having control over the generation of moves helps generate a better internal representation of the changing positions, an explanation that gains in plausibility given that Gobet and Retschitzki (1991) found that the level of controllability affects performance in chess. Another piece of evidence consistent with this explanation comes from a study by Wagner and Scurrah's (1971), in which a player showed a larger depth of search when he was pondering a position taken from his own game compared to unfamiliar positions. In the present study, as well as in Campitelli and Gobet's (2004) study, the moves were provided to the players who did not have this control. This may have caused them difficulties in generating internal representations, and thus forced even strong players to fixate on the board. One way to study this issue would be to compare players' performance in a similar task when they play their own game, as opposed to following a game played by others, as happened in the present experiment.

Finally, a topic to which we did not pay attention when we generated our hypotheses, but which is of interest, is the increase in performance over the three games played that happened in all players in both conditions. This result is consistent with Chase and Simon's (1973b) memory task in which the chess pieces were replaced by letters and performance improved with practice.

Conclusion

Using a multiple case study, we investigated two everyday psychological phenomena—averting gaze from the environment when generating internal representations and looking at parts of the environment to aid our internal cognitive processes—by combining two powerful research tools in psychology: recording of eye movements and blindfold chess. We suggested that reducing or increasing our connection with the environment is a function of the complexity of the task and the level of expertise of the individual performing the task.

We showed that all players chose to interact with the environment (supporting our hypothesis 2) at the same frequency (against our hypothesis 3) to aid their imagery processes. Moreover, the performance was a function of the skill level in the chess condition but not in the shape condition (supporting our hypothesis 1). We also provided additional data showing that players performed as predicted by their skill level in domain-specific tasks but not in general cognitive tasks. Together, our results showed that performance and, to some extent, interaction with the environment are a function of the expertise level and the complexity of the task. Moreover, our results are consistent with pattern-recognition theories that state that differences in performance between individuals with different levels of expertise in a specific domain are mainly due to the existence of domain-specific long-term memory patterns (Chase & Simon, 1973a; Gobet & Simon, 1996a), and not to general cognitive skills.

The result that the number of glances at the board was not a function of skill level is surprising, because strong masters are able to maintain internal representations of positions (e.g., when playing blindfold chess). In addition to the possibility that players' lack of control over the moves may have created difficulties in generating internal representations, we suggested that strong players might have been capable of performing the task without looking at the board, but that they chose to do so. The question as to whether this behaviour was a question of choice or necessity can be answered experimentally, and we proposed experiments to address this question.

References

Bachmann, T., & Oit, M. (1992). Stroop-like interference in chess players' imagery: An unexplored possibility to be revealed by the adapted moving-spot task. *Psychological Research, 54*, 27-31.

Binet, A. (1894). *Psychologie des grands calculateurs et joueurs d'échecs*. Paris: Hachette. [Reedited by Slatkine Ressources, Paris, 1981.]

Brandt, S. A., & Stark, L. W. (1997). Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience, 9*, 27-38.

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.

Campitelli, G., & Gobet, F. (2005). The mind's eye in blindfold chess. *European Journal of Cognitive Psychology, 17,* 23-45.

Charness, N. (1976). Memory for chess positions: Resistance to interference. *Journal of Experimental Psychology: Human Learning and Memory, 2*, 641-653.

Charness, N. (1981). Search in chess: Age and skill differences. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 467-476.

Charness, N. (1991). Expertise in chess: The balance between knowledge and search. In K. A. Ericsson & J. Smith (Eds.), *Studies of expertise : Prospects and limits* (pp. 39-63). Cambridge: Cambridge University Press.

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

Charness, N., Krampe, R., & Mayr, U. (1996). The role of practice and coaching in entrepreneurial skill domains: An international comparison of life-span chess skill acquisition. In K. A. Ericsson (Ed.), *The road to excellence* (pp. 51-80). Mahwah, NJ: Lawrence Erlbaum.

Charness, N., Tuffiash, M., Krampe R., Reingold, E., & Vasyukova, E. (2005). The role of deliberate practice in chess expertise. *Applied Cognitive Psychology, 19*, 151-165.

Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. *The Psychology of Learning and Motivation, 16*, 1-58.

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

Chase, W. G., & Simon, H. A. (1973b). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press.

Church, R. M., & Church, K. W. (1977). Plans, goals, and search strategies for the selection of a move in chess. In P. W. Frey (Ed.), *Chess skill in man and machine* (pp. 131–156). New York: Springer Verlag.

De Groot, A. D. (1946/1978). *Thought and choice in chess* (2 ed.). The Hague: Mouton.

De Groot, A. D., & Gobet, F. (1996). *Perception and memory in chess*. Assen: Van Gorcum.

Ericsson, K. A., & Staszewski, J. J. (1989). Skilled memory and expertise: Mechanisms of exceptional performance. In D. Klahr & K. Kotovski (Eds.), *Complex information processing: The impact of Herbert A. Simon.* Hillsdale, NJ: Erlbaum.

Fine, R. (1965). The psychology of blindfold chess. An introspective account. *Acta Psychologica, 24*, 352-370.

Glenberg, A. M., Schroeder, J. L., & Robertson, D. A. (1998). Averting gaze disengages the environment and facilitates remembering. *Memory & Cognition, 26*, 651-658.

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

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

Gobet, F., & Campitelli, G. (2007). The role of domain-specific practice, handedness and starting age in chess. *Developmental Psychology, 43,* 159-172.

Gobet, F., & Clarkson, G. (2004). Chunks in expert memory: Evidence for the magical number four… or is it two? *Memory*, 12, 732-747.

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

Gobet, F., & Retschitzki, J. (1991). Détresse apprise et jeu d'échecs: Rôle de la familiarité et de la similitude des tâches. *Revue Suisse de Psychologie, 20*, 97-110.

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 rapidly presented random chess positions is a function of skill. *Psychonomic Bulletin & Review, 2*, 159-163.

Gobet, F., & Simon, H. A. (1998). Expert chess memory: Revisiting the chunking hypothesis. *Memory, 6*, 225-255.

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

Gruber, H. (1991). *Qualitative Aspekte von Expertise im Schach.* Aachen: Feenschach.

Hebb, D. O. (1968). Concerning imagery. *Psychological Review, 76*, 466-477.

Holding, D. H. (1985). *The psychology of chess skill*. New Jersey: Lawrence Erlbaum.

Laeng, B., & Teodorescu, D. S. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science, 26*, 207-231.

Livshitz, A. (1988). *Test your chess IQ*. Oxford, UK: Pergamon Press.

Masunaga, H., & Horn, J. (2000). Characterizing mature human intelligence expertise development. *Learning and Individual Differences, 12*, 5-33.

Milojkovic, J. D. (1982). Chess imagery in novice and master. *Journal of Mental Imagery, 6*, 125-144.

Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood squares: Looking at things that aren’t there anymore. *Cognition, 76*, 269-295.

Saariluoma, P. (1991). Aspects of skilled imagery in blindfold chess. *Acta Psychologica, 77*, 65-89.

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

Saariluoma, P., & Kalakoski, V. (1997). Skilled imagery and long-term working memory. *American Journal of Psychology, 110*, 177-201.

Saariluoma, P., & Kalakoski, V. (1998). Apperception and imagery in blindfold chess. *Memory, 6*, 67-90.

Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research, 65*, 235-241.

Wagner, D. A., & Scurrah, M. J. (1971). Some characteristics of human problem-solving in chess. *Cognitive Psychology, 2*, 454-478.

Waters, A. J., Gobet, F., & Leyden, G. (2002). Visuo-spatial abilities in chess players. *British Journal of Psychology, 30*, 303-311.

Appendix

Example of a sequence of moves and tests as used in the main experiment of this paper. The position at the top of the Figure remained unchanged and in view for the entire sequence of moves. The 1st column indicates the move number, the 2nd column indicates moves for white, and the 3rd column indicates moves for black. The fourth column shows the tests carried out during the game. Throughout, the algebraic chess notation was used (see footnote 5).

INSERT FIGURE A1 ABOUT HERE

Table 1.

Percentage correct in each game and correlation with skill level

	Chess			Shape		
	game 1	game 2	game 3	game 1	game 2	game 3
Performance						
GM	66.7	86.1	94.4	16.7	19.4	61.1
IM	50.0	66.7	80.6	0	16.7	52.8
CM	25.0	41.7	55.6	12.5	16.7	27.8
CB	16.7	16.7	36.1	8.3	22.2	30.6
corr. skill x performance	.94	.98	.98	-.01	-.58	.88
p (one-tailed)	<.05*	<.05*	<.05*	>.1	>.1	>.06

Note. * significant correlation using p < .05 as criterion. For the skill x performance correlation the variable Elo was used (GM = 2550, IM = 2500, CM = 2100 and CB = 1750).

Table 2.

Relative numbers of vertical and horizontal shifts together with correlations with skill and performance

	Chess			Shape		
	game 1	game 2	game 3	game 1	game 2	game 3
Vertical shifts						
GM	0.92	1.21	1.33	1.87	1.89	1.82
IM	1.68	1.89	1.55	2.27	2.14	1.89
CM	1.35	1.49	1.60	1.45	1.85	1.59
CB	1.27	0.97	1.32	0.60	0.74	0.87
corr. Skill x vertical shifts	-.03	.60	.19	.95	.88	.95
p (one-tailed)	> .1	>.1	> .1	< .05*	> .05	< .05*
corr. performance x vertical shifts	-.30	.42	.03	-.24	-.89	.70
p (one-tailed)	> .1	> .1	> .1	> .1	> .05	> .1
Horizontal shifts						
GM	1.73	1.47	0.99	1.60	1.90	1.22
IM	1.77	1.69	2.11	2.29	2.24	2.35
CM	1.56	1.57	1.58	1.56	1.32	1.21
CB	1.21	1.44	1.57	1.67	1.51	1.68
corr. Skill x horizontal shifts	.97	.50	-.08	.42	.76	.16
p (one-tailed)	< .05*	> .1	> .1	> .1	> .1	> .1
corr. performance x horizontal shifts	.84	.30	-.24	-.91	-.26	.20
p (one-tailed)	> .07	> .1	> .1	< .05*	> .1	> .1

Note. * significant correlation using p < .05 as criterion.

Table 3.

Additional data: Chess-related tasks

	GM	IM	CM	CB	corr. with skill	p (1-tailed)
Quick problem solving						
Performance (%)	46.9	36.9	8.2	4.1	.94	<.05*
Long-term memory						
Session 1 (%)	93	92	59	58	.93	<.05*
Session 2 (%)	87	80	60	54	.96	<.05*
Session 3 (%)	82	69	56	50	.93	<.05*
Session 4 (%)	67	63	60	50	.96	<.05*
Short-term memory						
Standard positions (%)	89.2	88.5	48.3	43.7	.95	<.05*
Quiet positions (%)	89.4	95.2	69.2	59.6	.97	<.05*
Complex positions (%)	68.8	65.9	44.2	36.3	.98	<.05*
Practice						
Cumulative hours studying	6890	7904	1872	416	.96	<.05*
Cumulative hours playing	7722	7072	2704	1326	.98	<.05*

Table 4.

Additional data: Non-chess-related tasks

	GM	IM	CM	CB	corr. with skill	p (1-tailed)
Reaction time						
Session 1 (ms)	376	349	417	370	-.30	>.1
Session 2 (ms)	315	392	356	351	-.03	>.1
Session 3 (ms)	344	371	339	315	.84	>.07
Session 4 (ms)	323	362	317	332	.32	>.1
Long-term memory						
Session 1 (%)	91.6	86.3	86.3	89.5	.07	>.1
Session 2 (%)	93.5	75.0	86.9	78.2	.32	>.1
Session 3 (%)	87.5	73.9	84.4	71.9	.47	>.1
Session 4 (%)	78.7	71.3	67.0	71.3	.53	>.1
Memory span						
Letters	6.5	5.5	6.0	7.0	-.62	>.1
Digits	6.0	8.0	7.5	6.5	.12	>.1
Spatial	4.0	4.0	4.0	4.5	-.84	>.07
Short-term memory						
Random (%)	26.9	20.9	18.9	16.9	.84	>.07
Shape-random (%)	13.3	19.0	14.3	10.5	.69	>.1
Shape-game (%)	18.3	11.5	15.4	19.2	-.49	>.1

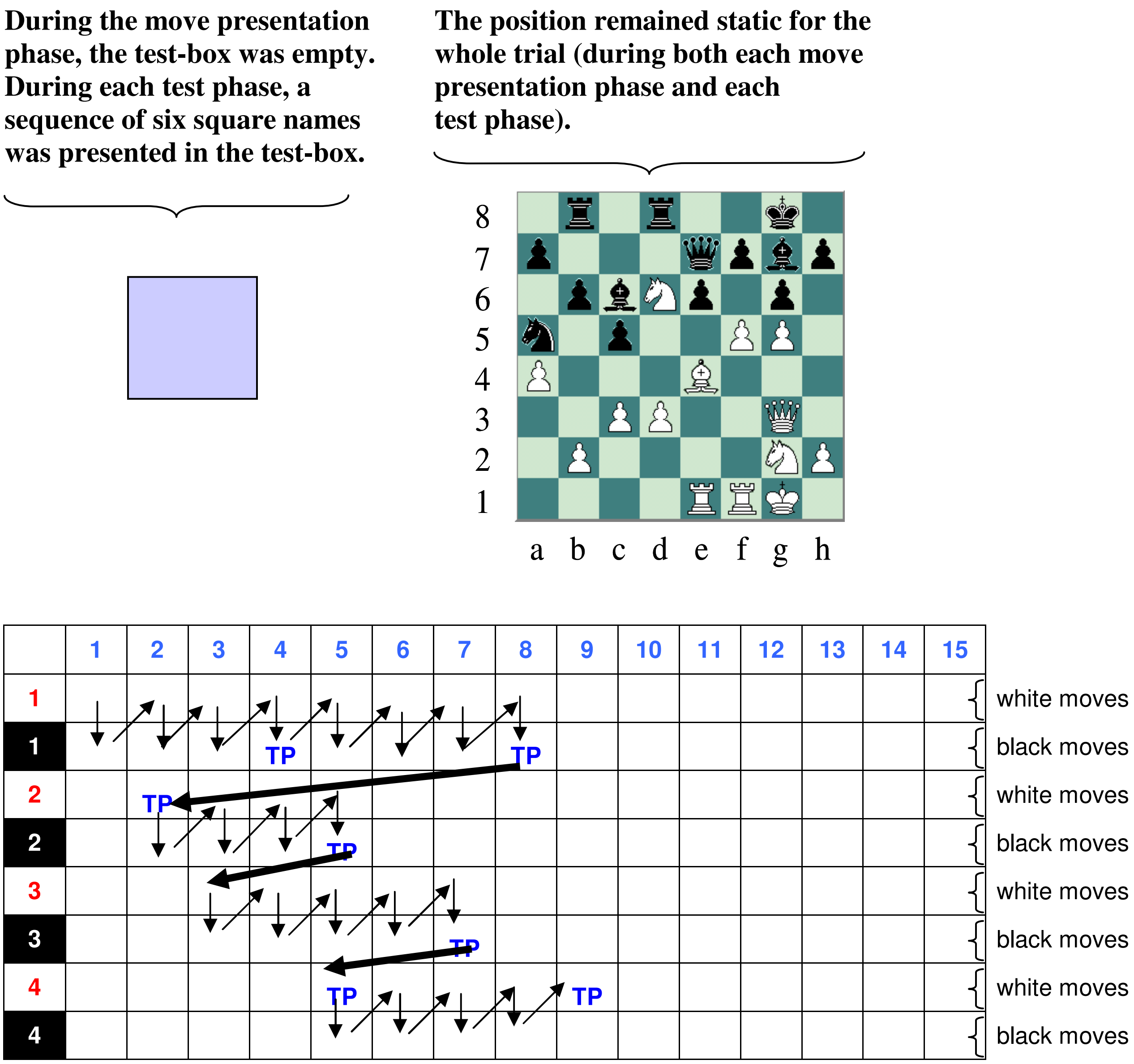
Note. * significant correlation using p < .05 as criterion.

Figure captions

Figure 1. Illustration of a typical game presentation in the chess condition (the same procedure was used for the shape condition). All the moves were presented in algebraic notation. White moves were presented in the rows 1 to 4 with red (grey in the figure) number in a white background and black moves were presented in rows 1 to 4 with white numbers in a black background. The small arrows illustrate the sequence of moves in the move presentation phase (note that these arrows were not presented in the actual trials). Every white move was followed by a black move, and then a new white move was presented, and so on. At some point of the sequence, no moves were presented, and a test phased (TP) started. The cue for the participants that a test phase had started was that the grid became blue (this is not shown in this figure; see Figure 2). Note that the label "TP" was not presented during the trial; it is used here to indicate how the test and move presentation phases were intermingled. Moreover, the coordinates "a" to "h" and 1 to 8 were not presented during the trial. The big backward arrows indicated to the players that they had to recover the position at a previous stage of the sequence of moves that had been already presented. In the example, the first big arrow indicates that the players, after being shown eight white and black moves, had to recover the position after the second white move. After this, a new TP started, and then, a new sequence of moves was presented.

Figure 2. Test session (shape condition). The grid became blue (grey in the figure) and the test box showed the name of a board square. The players had to say out loud which piece was located in that square. Note that the coordinates "a" to "h" and 1 to 8 were not presented during the trial.

Figure A1. Example of a sequence of moves and tests.

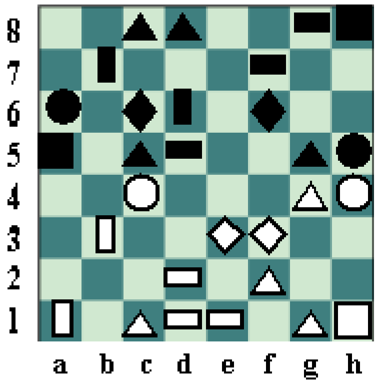


During the move presentation phase, the test-box was empty. During each test phase, a sequence of six square names was presented in the test-box.

The position remained static for the whole trial (during both each move presentation phase and each test phase).

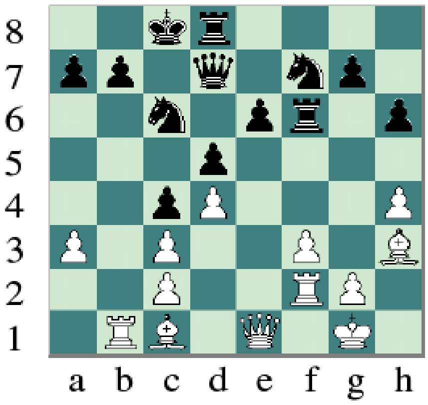
Figure 1

C7



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1															
1															
2															
2															
3															
3															
4															
4															

Figure 2



1.	♗c1-f4	♙g7−g5	
2.	♖f2−e2	♖d8−e8	
3.	♗f4−h2	♙g5−h4	
4.	♕e1−h4	♘f7−g5	
5.	♗h3−g4	♕d7−g7	
6.	♖b1-e1	♙b7−b6	
7.	♕h4−h5	♕g7−d7	
8.	♙f3−f4	♘g5−e4	
9.	♖e2−e4	♙d5−e4	TEST: d7, f4, g5, f6, g4, e1
10.	♙d4−d5	♘c6−d8	
11.	♕h5−e5	♖f6−f5	
12.	♙d5−e6	♕d7−d2	
13.	♕e5−e4	♖f5−d5	
14.	♙e6−e7	♔c8−c7	
15.	♙f4−f5		TEST: d2, e5, d7, d5, g4, c6
Back to position after 12 ... ♕d7−d2			TEST: e6, e4, h5, f5, d5, c6
Back to position after 6 ... ♙b7−b6			TEST: g7, g5, f4, h4, e4, d5

Figure A1

Footnotes

[1] There is some disagreement as to whether seeing an empty board while playing (simultaneous) blindfold chess is of any help. Fine (1965) suggested that it was an hindrance, while Koltanowski, who held the world record in simultaneous blindfold chess, thinks it helps (see Campitelli & Gobet, 2005).

[2] Elo (1978) developed the rating scale that is now used by the World Chess Federation (FIDE). The scale has a normal distribution and a standard deviation of 200 points. The best player of the world has around 2800 points and the weakest less than 1200. In the psychology literature, players between 1600 and 1800 are called Class B, between 1800 and 2000 Class A, between 2000 and 2200 Experts, and players with more than 2200 are considered masters. FIDE awards players with titles; players above 2300, 2400 and 2500 are usually called FIDE masters, international masters, and international grandmasters, respectively. In this article we use “candidate master” to denote players between 2000 and 2200 Elo points and “expert” as a general word for individuals possessing expertise.

[3] In the algebraic notation all the squares of the board have a name. This name is formed by a letter followed by a number. The letter corresponds to the column, and the number indicates the row (see chessboard in Figures 1 and 2). The move is indicated by including the symbol or the initial letter of the name of the moving piece (e.g., R for Rook), followed by the name of the origin square (e.g., b2) and then the name of the destination square (e.g., b7). All the participants in this study were familiar with the algebraic notation.