

Mental imagery and chunks: Empirical and computational findings

ANDREW J. WATERS

University of Texas M. D. Anderson Cancer Center, Houston, Texas

AND

FERNAND GOBET

Brunel University, Uxbridge, England

To investigate experts' imagery in chess, players were required to recall briefly presented positions in which pieces were placed on the intersections between squares (intersection positions). Position types ranged from game positions to positions in which both the piece distribution and the location were randomized. Simulations were run with the CHREST model (Gobet & Simon, 2000). The simulations assumed that pieces had to be centered back, one by one, to the middle of the squares in the mind's eye before chunks could be recognized. Consistent with CHREST's predictions, chess players ($N = 36$), ranging from weak amateurs to grandmasters, exhibited much poorer recall for intersection positions than for standard positions (pieces placed on the centers of the squares). For the intersection positions, the skill difference in recall was larger for game positions than for the randomized positions. The participants recalled bishops better than they recalled knights, suggesting that Stroop-like interference impairs recall of the latter. The data supported both the time parameter in CHREST for shifting pieces in the mind's eye (125 msec per piece) and the seriality assumption. In general, the study reinforces the plausibility of CHREST as a model of cognition.

Ever since Binet's (1894/1981) work on blindfold chess, psychologists have investigated the role of mental images in expert problem solving (e.g., Larkin, McDermott, Simon, & Simon, 1980; Larkin & Simon, 1987; Paige & Simon, 1966; Simon, 1978; Tabachneck-Schijf, Leonardo, & Simon, 1997). The concept of chunking has also been shown to be essential in experts' perception, memory, and problem solving (e.g., Campitelli & Gobet, 2005; Chase & Simon, 1973a, 1973b; de Groot, 1946/1978; Kalakoski, 2006; Saariluoma, 1995). However, despite the wealth of research in which chunking and mental imagery have been investigated as separate topics, less is known about the interaction between chunking and mental imagery. The goal of this article is to provide experimental evidence about this interaction and to test the predictions of a well-established computational theory of expertise.

Mental Imagery in Chess

Chase and Simon's (1973b) influential chunking theory proposed that pattern recognition explains (1) how experts show a remarkable memory for domain-specific material and (2) how search can be carried out in the mind's eye, in which future positions are imagined. According to these authors, the mental processes used in chess playing are similar to those identified in mental rotation and other imagery tasks (Shepard & Cooper, 1982). Chase

and Simon's (1973b) ideas about the mind's eye have not been thoroughly tested, but the existence of chunks is well established (Chase & Simon, 1973b; Gobet & Clarkson, 2004; Gobet & Simon, 1998).

Empirical evidence on imagery in chess comes from studies on blindfold chess (e.g., Saariluoma & Kalakoski, 1997) and from studies in which attempts have been made to measure the time needed to move a piece in the mind's eye (Church & Church, 1977; Gruber, 1991; Milojkovic, 1982; Waghorn, 1988). In general, these experiments have confirmed that (1) chess memory has a significant visuospatial component and (2) mental imagery has an important role in chess. For example, Bachmann and Oit (1992) studied mental imagery using a variation of Attneave and Curlee's (1983) moving spot task. Chess players and nonplayers were presented with either an 8×8 grid or a chessboard. They were then required to close their eyes, listen to a sequence of instructions about the moves of the spot or a chess piece (up, down, right, or left), and imagine following the spot or the piece as it moved. At the end of the sequence of moves, the participants had to indicate the end position of the spot or the piece. There were no skill differences in the moving spot (8×8 grid) condition, but nonplayers made more errors than did chess players in the moving chess piece (chessboard) condition. Furthermore, in the latter condi-

F. Gobet, fernand.gobet@brunel.ac.uk

tion, skilled players tended to show Stroop-like interference when required to mentally shift a piece in an atypical fashion. For example, chess players found it difficult to imagine a bishop moving horizontally (which is incongruent with its typical diagonal movement).

As was noted above, mental imagery played an essential role in Chase and Simon's (1973a, 1973b) influential chunking theory. 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 & Simon, 1996b). The template theory has led to two related computational implementations: one whose aim is to simulate search behavior (SEARCH; Gobet, 1997), and another whose predominant aim is to simulate perception, learning, and memory (CHREST; de Groot & Gobet, 1996; Gobet et al., 2001; Gobet & Simon, 2000; Gobet & Waters, 2003).

CHREST

CHREST is a model of learning and expertise that has accounted for data on perception, learning, and memory in chess (de Groot & Gobet, 1996; Gobet et al., 2001; Gobet & Simon, 2000; Gobet & Waters, 2003), the use of diagrammatic information in physics (Lane, Cheng, & Gobet, 2000), the acquisition of vocabulary (Jones, Gobet, & Pine, 2005), and the acquisition of syntactic structures (Freudenthal, Pine, & Gobet, 2005, 2006). CHREST, written in Common Lisp, can be obtained from the second author.

CHREST consists of four main components: a simulated eye; a long-term memory (LTM), in which chunks

are stored; a visual short-term memory (STM) with a capacity of three items; and a mind's eye system (see Figure 1 for an overview). Briefly, CHREST is a self-organizing, dynamical system, in which chunks are accessed by traversing a discrimination net (see de Groot & Gobet, 1996, for details). A discrimination net is a tree-like structure consisting of a set of nodes (chunks) connected by links. The links have tests, which are applied to check features of the external stimuli.

There has been debate as to how chunks are coded in chess. Some authors (e.g., Simon & Gilmarin, 1973) have proposed that chunks encode information about location (localization assumption). Others (e.g., Holding, 1985) have suggested that chunks do not encode location information, but only the pattern of relation between pieces. Thus, the same chunk could be used for coding a pattern of pieces at the bottom left of the board or the top right of the board. Experiments in which boards were modified by translation (Saariluoma, 1994) or by a mirror image (Gobet & Simon, 1996a) have supported the localization assumption.

Chunks that are often recognized evolve into more complex data structures, known as *templates*, which have *slots* allowing variables to be instantiated rapidly (filling in information into a template slot takes 250 msec). In particular, information about piece location, piece type, or chunks can be (recursively) encoded into template slots. Slots are created at chunks where there is substantial variation in squares, pieces, or groups of pieces in the test links below. In addition to slots, templates contain a *core*, basically similar to the information stored in chunks. Chunks and templates can be linked to other information stored in

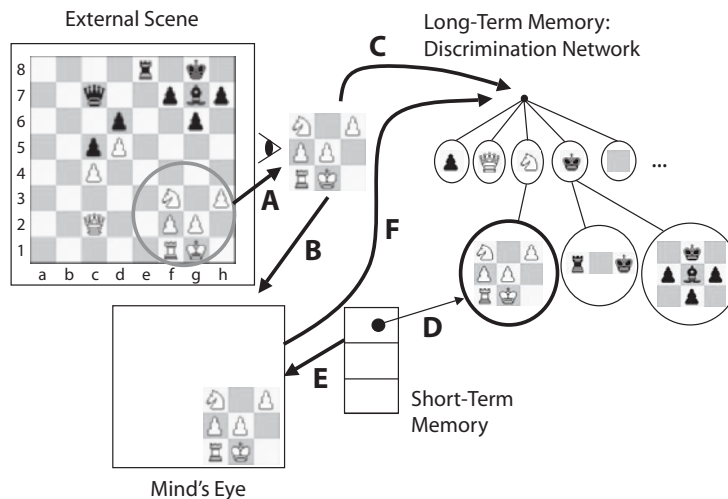


Figure 1. CHREST consists of four components: a simulated eye, a discrimination network giving access to long-term memory (LTM), a short-term memory (STM), and a mind's eye. The simulated eye selects a portion of the external display (the *visual field*, shown by A in the figure). This information is sent to both the mind's eye (B) and LTM (C). If the information is recognized in LTM by accessing a node (i.e., a chunk), a pointer to this chunk is put in STM (D), and the information is unpacked in the mind's eye (E). In turn, the information in the mind's eye can be used to access a node in LTM (F). Note that LTM chunks encode information about location (e.g., the two white pieces on the first row would be encoded as a rook on f1 and a king on g1).

LTM, such as moves, plans, and tactical motives (see Ferrari, Didierjean, & Marmèche, 2006, for data supporting the idea that chunks are associated with possible moves).

Eye movements are directed by a combination of acquired knowledge, mediated by the structure of the discrimination net, and heuristics (e.g., heeding a square attacked by the piece located on the currently fixated square; see de Groot & Gobet, 1996, for details).

The Mind's Eye in CHREST

The mind's eye construct in CHREST is similar to Kosslyn's (1994) visual buffer and Baddeley's visuosketchpad (Baddeley, 1986; Logie, 1986). It is specified in less detail than the former—in particular, with respect to anatomical and neural considerations—but with more detail than the latter, which lacks a computational implementation. The architectural assumptions in CHREST resemble those in other models developed by Simon—in particular, CaMeRa (Tabachneck-Schijf et al., 1997) and EPAM IV (Richman, Staszewski, & Simon, 1995)—and can be traced back to Chase and Simon's (1973a) construct of the mind's eye.

The mind's eye stores perceptual structures, both from external inputs and from memory stores, for a short time. The visuospatial information stored there can be subjected to visuospatial mental operations. In the mind's eye, the internal representation of the external scene is encoded as a network of nodes and links (e.g., Larkin & Simon, 1987; Newell & Simon, 1972). Note that the information in the mind's eye abstracts much from perceptual information impinging the retina, which makes the task of pattern recognition mechanisms easier than with external perceptual information.

A recurring feature of CHREST, influenced by Simon's earlier work (e.g., Simon, 1969), is the emphasis on cognitive limitations. There are limitations in memory capacity (visual STM can hold only three items) and learning rates (it takes about 8 sec to create a new chunk). Another limitation is that information in the mind's eye decays rapidly, within around 250 msec (Averbach & Coriell, 1961; Kosslyn, 1994).

In addition, CHREST makes several assumptions about the processes that are carried out in the mind's eye. For chess, these processes include the time to move a piece mentally; for problem solving in physics, these processes include instructions for drawing lines or more complex geometric figures (Lane et al., 2000). These mental processes are assumed (1) to take a definite amount of time (see below) and (2) to be carried out serially (*seriality assumption*; see Kosslyn, Cave, Provost, & von Gierke, 1988, for data supporting the assumption that mental images are generated serially). In addition, CHREST includes mechanisms linking LTM, STM, and the mind's eye. It is assumed that learning leads to the creation of chunks in LTM. When a chunk is elicited, either by external or internal information, a pointer to it is placed in STM.¹ Concurrently, the visuospatial information referred to in LTM by this pointer is unpacked in the mind's eye. Since information in the mind's eye fades rapidly, it needs to be refreshed regularly.

Overview of the Study

Several assumptions behind CHREST have been directly tested (de Groot & Gobet, 1996; Gobet & Jackson, 2002; Gobet & Simon, 2000). However, there have been no experimental tests of CHREST's assumptions and mechanisms about imagery. Using the few experiments available (noted above), de Groot and Gobet (1996, p. 236) proposed two parameters for the time to move pieces in the mind's eye: The *base* parameter refers to the time needed to start the process of generating a move, and the *square* parameter estimates the time needed to move a piece over one square in the mind's eye. With players who are not novices, the base parameter was set to 100 msec, and the square parameter was set to 50 msec. The present study directly tested the validity of these parameters.

In addition, a strict interpretation of the localization assumption implies that decoupling pieces from their natural ground (the chessboard)—for example, by shifting them diagonally—should make access to chunks harder, because the pieces would have to be individually shifted back to the center of their squares in the mind's eye to enable access to chunks. An alternative prediction would be that the relations between pieces are identifiable without reference to the ground (the board) and, thus, no shifting back is necessary. To decouple figure and ground, we created positions in which the pieces were placed at the intersection of squares (rather than being placed in the middle of the squares). We will call these stimuli *intersection positions*. If chunks are recognized without the need to recenter pieces, recall for the intersection positions should not differ from that for the standard positions. If, on the other hand, pieces need to be recentered before chunks can be recognized, there should be a decrease in performance. The size of this decrease can be predicted by running simulations with CHREST.

Two additional variables were manipulated. First, positions with different levels of structure were used (from game position to fully randomized positions). This manipulated the ease with which chunks could be accessed in LTM. Second, different skill levels were used, which was assumed to control for the number of chunks held in LTM.

In the remainder of the article, we first will present a simulation that assesses recall for different position types across different skill levels (net sizes) (Computer Simulation 1). We then will present human data that can be used to assess recall for the same position types (Human Study 1). This will be followed by an additional human study in which the specificity of the effects observed in Human Study 1 was assessed (Human Study 2) and an additional simulation in which the seriality assumption for the mind's eye was examined (Computer Simulation 2).

COMPUTER SIMULATION STUDY 1

Learning Phase

During learning, the program scans a large number of positions. For each position, the simulated eye is moved around the board, and patterns within CHREST's visual field (defined as plus or minus two squares from the fixation point) are sent as input to the discrimination net, where the learning

mechanisms of familiarization, discrimination, and template formation are applied (see de Groot & Gobet, 1996, and Gobet & Simon, 2000, for details on the model).

Testing Phase

Encoding. The model moves its simulated eye around the board and attempts to recognize chunks (or templates). All of the simulations reported below used a presentation time of 5 sec. We employed the same version of CHREST as that used by Gobet and Simon (2000) and Gobet and Waters (2003). For the simulation of the intersection positions, we extended the model by adding as few assumptions as possible related to the time needed to carry out operations in the mind's eye. The augmented model attempts to memorize the intersection positions by serially moving pieces (up to three) within the visual field to the center of the square in the mind's eye and then sorting the (shifted) pattern of pieces through the discrimination net.

Figure 2 illustrates the processes involved in shifting a group of three pieces in the mind's eye. It is assumed that the internal representations of the objects present in the external display (pieces on the intersection board) maintain a high activation as long as one is looking at these objects (cf. Kosslyn, 1994). By contrast, after shifting, the imaged (shifted) objects are subject to decay. Shifting the object or recognizing a chunk in the mind's eye resets the level of

activation of these objects to high. As with chunk recognition from an external input, recognition from the mind's eye leads to a pointer's being placed in STM (see above).

At time 0, the three pieces are perceived in the external display. At time 125 msec, the bishop has been shifted, and its activation is high. At time 250 msec, the rook has been shifted (high activation); in the meantime, the activation of the bishop has decayed to medium. At time 280 msec, a chunk (rook + bishop) has been recognized, a pointer to it has been entered into STM, and the activation of its components has been set to high. Sorting information in the discrimination network has a base cost of 10 msec, plus 10 msec per item, giving a total of 30 msec. At time 405 msec, the knight has been shifted, and the activation of the chunk (rook + bishop) has decreased to medium. At time 445 msec, a three-piece chunk (rook + bishop + knight) has been recognized, and a pointer to it has replaced the previous pointer in STM.

Recall. During recall, the model (re-)places pieces sequentially, using the information provided by the chunks in STM. Possible conflicts (e.g., two chunks propose that two different pieces have been placed on the same square) are resolved in sequence, on the basis of the frequency with which each placement is proposed. The pieces are always reshifted to the appropriate intersection (southeast [SE] corner). This reshifting assumption was predicated on the assumption that the direction of reshifting was not problematic for the humans. As will be noted below, participants learned the direction in which the pieces were shifted during practice, and the presence of pieces on the "back" white row (but not on the "back" black row) gave additional visual cues on every trial. In addition, the software assisted in piece re-placement (clicking a piece on a square at a location sufficiently close to the intersection placed the piece at the intersection) (this was not simulated in the model).

For the time needed to shift a piece (transition time), we used the parameters already in CHREST for the time needed to move a piece diagonally by players who were not novices (100 msec as base time and 50 msec for each square traversed). We therefore assumed that, to recenter the pieces in the intersection positions, it takes the base time plus half of the time to traverse a square diagonally—that is, 125 msec in total. This was true for discrimination nets of all sizes.

Method

Materials

In common with Gobet and Waters (2003), we used five position types. Game positions were taken from master games without any change. Random positions were constructed by randomly reassigning the pieces of a game position to new squares. In "truly" random positions, not only the location of the pieces was randomized, but also the distribution of pieces (e.g., there could be 12 white kings in a position, contrary to the standard chess rules). One-third and two-thirds truly random positions were positions for which one third and two thirds of the pieces were truly randomized.

The testing stimuli were created following the procedure described in Gobet and Waters (2003). Five hundred stimuli were selected, using random sampling without replacement from a database of 3,100 positions. These positions were taken from master-level games, after about 20 moves. In the game condition, the stimuli were

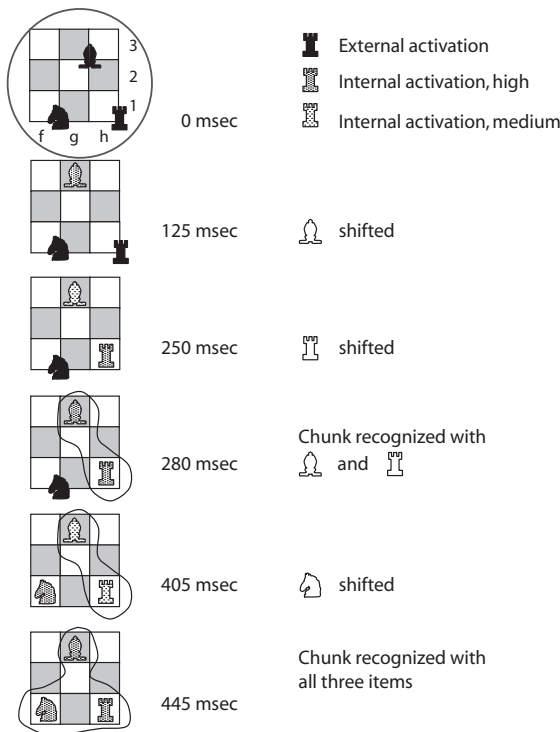


Figure 2. Processes involved in shifting pieces in the mind's eye in CHREST (see the text for details). The circle in the top diagram reminds the reader that the diagram is within the current visual field, and the coordinates remind the reader that chunks encode information about location.

kept unmodified. The algorithms described in Gobet and Waters were used to generate the four types of random positions. The same procedures were used to generate the intersection stimuli, except that the end product was manipulated by shifting all the pieces to the SE corners of the squares. Examples of all the position types for the intersection positions are shown in Figure 3A.

Procedure

To obtain quantitative predictions, we used the same nets as those chosen by Gobet and Waters (2003). The selection procedure was as follows. During learning, the program scanned a database of about 50,000 positions. The positions were middle-game positions, taken from master-level games played in the last 50 years. This resulted in an original pool of 16 nets, from which 4 nets (with 1,010, 3,008, 15,003, and 300,009 chunks) were chosen. These were selected because they most closely matched the mean recall of the four groups of human participants on standard game positions. To facilitate interpretation of the data analyses, the 300,009-chunk net was considered to have a rating of 235.6 in “human” units (the mean skill rating of our top group of humans; see Gobet & Waters, 2003, for further detail on this strategy).² The 15,003-chunk net was considered to have a rating of 201.2 in human units (the mean skill rating of our second group of humans), the 3,008-chunk net a rating of 150.9, and the 1,010-chunk net a rating of 112.3. The simulations below are based on the recall of 500 positions of each type. Each position was presented for a simulated time of 5 sec.

Data Reduction and Analysis

All statistical analysis was carried out using SAS software. To test differences in recall for the standard and the intersection positions, we used a two-way layout (standard vs. intersection) × position type (five levels: game, 1/3, 2/3, random, or 3/3) ANOVA. On the intersection positions, we used linear regression to predict recall performance (the dependent variable) from net size (the independent variable, expressed in human units). We did this for each position type (game, 1/3, 2/3, random, and 3/3) separately. We used the unstandardized parameter estimates from these models as an index of the net size effect and used the *p* value of the parameter estimate to determine whether the slope was significantly different from zero. To determine whether the net size effect for one position type was steeper than those for other position types, we used regression analysis (proc glm in SAS) to test for net size × position type interactions, where position type was entered as a repeated measures variable. To compare the net size effect for the game positions with the mean net size effect for the other positions, we used the helmert comparison option.

Results and Discussion

Table 1 shows recall as a function of net size and position type, for both standard (left side of table) and intersection (right side of table) positions; for the latter, we will focus, for the time being, on the serial transition type (upper row). Recall was clearly worse for the intersection positions. A two-way layout (standard vs. intersection) × position type (five levels) ANOVA yielded a main effect of layout [$F(1,3) = 29.9, p < .05$] and a main effect of position type [$F(4,12) = 40.3, p < .0001$]. There was also a layout × position type interaction [$F(4,12) = 15.8, p < .0001$], indicating that deterioration in recall for the intersection positions was moderated by position type. An inspection of Table 1 indicates that the deterioration in recall for the intersection positions (vs. standard positions) was most pronounced for the game positions, because the model is slowed down by having to carry out the mental transformations. This prevents it from sampling a sufficient number of squares to access large chunks.

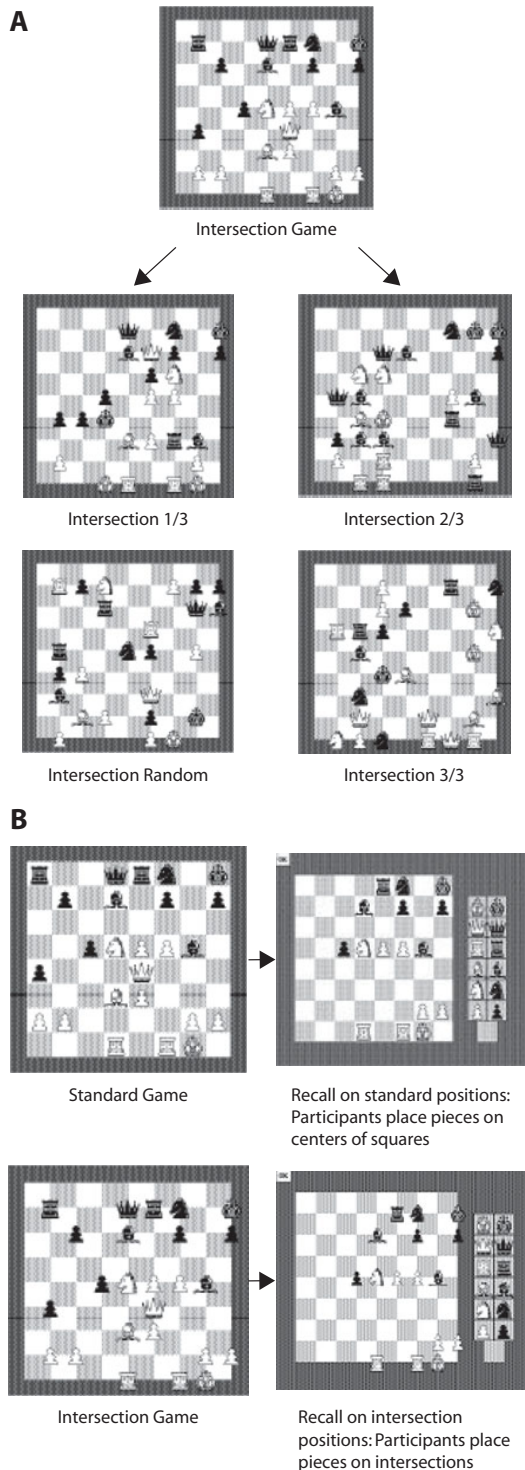


Figure 3. (A) Examples of the five position types used in the experiment (intersection positions). (B) Recall requirements for standard (top) and intersection (bottom) positions. For intersection positions, the participants were required to place the pieces on the intersections of the squares.

Table 1
Recall by Net Size/Skill and Position Type

Net Size	Skill	Standard					Trans.	Intersection					R^2	AAD	SSE
		Game	1/3	2/3	Random	3/3		Game	1/3	2/3	Random	3/3			
300K		79.5	52.8	26.2	21.7	15.2	serial	44.8	27.9	17.5	15.3	10.7	0.94	5.1	165.7
	GMs	83.7	28.7	20.3	18.3	14.8	parallel	68.9	42.2	24.1	19.8	14.2	0.95	15.5	1,608.5
15K		62.5	36.3	19.5	16.6	11.3	serial	37.9	22.9	14.5	11.4	7.5	0.94	2.6	42.2
	Experts	61.3	25.9	19.1	17.0	16.3	parallel	54.3	31.7	18.6	15.3	10.7	0.96	8.3	549.0
3K		49.4	28.1	17.1	14.4	8.7	serial	30.4	19.4	12.2	9.5	5.4	0.91	2.7	44.4
	Class A/B	55.1	22.2	16.2	14.1	13.7	parallel	47.5	26.4	16.4	13.7	8.0	0.95	6.7	341.3
1K		37.6	24.5	15.9	13.3	7.2	serial	23.7	15.7	10.4	8.1	4.1	0.90	2.2	30.5
	Class C/D	40.1	16.9	15.4	12.3	12.0	parallel	33.6	23.0	15.4	12.4	6.1	0.88	5.4	182.4

Note—Mean percentage of recall by net size/skill and position type. Data are shown for transition time = 125 msec. The human data, broken down by skill level, are shown for comparison (data in bold). Trans, transition type; serial, pieces shifted in series (one by one) in the mind's eye; parallel, pieces shifted in parallel (as a group) in the mind's eye; AAD , average absolute deviation (the smaller, the better); SSE , sum of squared errors (the smaller, the better). Grandmasters (GMs), $n = 7$; experts, $n = 12$; Class A/B, $n = 10$; Class C/D, $n = 7$.

For the intersection positions (right side of Table 1), the data indicate that the effect of net size is more robust in the game condition than for the randomized positions. The 300K net achieved 44.8% recall of the game positions, whereas the 1K net achieved only 23.7% recall. In contrast, for the 3/3 positions, the difference in recall was much smaller (10.7% vs. 4.1% for the 300K and 1K nets, respectively). Numerical estimates of the net size effects were the following: game = 0.173, 1/3 = 0.092, 2/3 = 0.050, random = 0.051, and 3/3 = 0.057. All of the net size effects were significantly different from zero ($p < .05$), except for those for the random positions, where the effect was a trend ($p = .058$). Statistical comparison of the net size effect for the game, 1/3, 2/3, random, and 3/3 position types indicated a robust net size \times position type interaction [$F(4,8) = 66.1, p < .001$]. The effect of net size was significantly larger for the game positions than for the mean of the other position [$F(1,2) = 571, p < .005$].

In summary, four CHREST nets were created, corresponding to four levels of chess skill. The program then recalled briefly presented positions (either standard or intersection) that ranged from game positions to fully random positions. Recall was poorer for the intersection positions than for the standard positions, particularly for the game positions. For the intersection positions, the skill effect was larger for the game positions than for the randomized positions.

HUMAN STUDY 1

As was noted earlier, this experiment was part of a larger study in which standard positions were also presented. The main focus in Gobet and Waters (2003) was to test the conflicting predictions of CHREST and the constraint attunement hypothesis (Vicente & Wang, 1998) on the role of constraints in expert memory. In the present study, the aims were as follows. First, we wanted to test CHREST's predictions on recall for the intersection positions. Thus, we examined whether (1) recall for the intersection positions was impaired (vs. recall for standard positions), (2) the effects of skill were larger for the game positions

than for the randomized positions, and (3) the effects of skill were significant for each position type.

Second, given that the intersection task likely involves imagery, we also examined whether we could detect Stroop-like interference effects that had been previously observed using a different imagery task (Bachmann & Oit, 1992). To do so, we compared recall of bishops and knights. These two pieces are considered to be of equal value (3 points) and, therefore, are matched in terms of "economic" salience on the chessboard. (It would be difficult to compare recall of rooks and bishops, because the former have greater value and might attract focus for this reason.) Critically, the bishop moves along a diagonal line, and so mental imagery on the intersection task is congruent with its typical movement. In contrast, the knight does not move diagonally, and so mental imagery with this piece is less congruent with its typical movement.³ Thus, if mental imagery were involved, we would expect participants to recall bishops better than they recall knights (Bachmann & Oit, 1992).

Method

Participants

Thirty-six participants (mean age = 28.6 years, $SD = 8.0$) completed the study; they were the same participants as those who took part in the study reported by Gobet and Waters (2003), and the data reported here was obtained from the 2003 study. The top group (*grandmasters*, $n = 7$) were players with British Chess Federation (BCF) ratings above 225. The second group (*masters/experts*, $n = 12$) had BCF ratings between 175 and 224. This group contained 3 international masters, 1 FIDE master, 1 female grandmaster, 1 female international master, and 6 experts. The third group (*Class A/B players*, $n = 10$) consisted of players with ratings between 125 and 174; these players are considered moderate to strong club players. The final group (*Class C/D players*, $n = 7$) contained players with ratings less than BCF 125; although these players are considered weak club players, they are far from being novices. Further details of the participants and experimental procedures are available in Gobet and Waters.

Materials

Chess stimuli. The same five types of positions were used as in the simulation study (see Figure 3A). Twenty-five positions (with

an average of 25 pieces) were taken from master games after about 20 moves and were randomly assigned to one of the five types of intersection positions for each player. There were thus five positions in each condition.

Presentation software and hardware. Chess stimuli were presented on a portable Apple Macintosh computer using specialized software for presenting chess stimuli and recording responses (see Gobet & Simon, 1998, for a detailed description of the software used). The participants were required to use a mouse to select pieces, move pieces onto squares (standard positions) or intersections between squares (intersection positions), and delete pieces (Figure 3B). To go on to the next trial, the participant pressed an OK button on the top left corner of the computer screen.

Visual memory test. All the participants completed a test of visual memory (VM; the Shape Memory Test [MV-1] of the Educational Testing Service Kit of Factor-Referenced Cognitive Tests; Ekstrom, French, Harman, & Derman, 1976).

Procedure

The participants completed the VM test, followed by the recall task. On each trial, a position was presented for 5 sec. The screen then was blank for 2 sec, and then an empty chess board appeared. The participants were instructed to try to recall the positions as completely and as accurately as possible. For standard positions, the pieces were presented in the center of the squares, and the participants were instructed to replace the pieces in the center of the squares (Figure 3B, upper panel). For intersection positions, the pieces were presented at the intersection of the squares (always in the SE direction), and the participants were instructed to replace the pieces at the intersection of the squares (Figure 3B, lower panel). In this condition, the software allowed the pieces at the intersection of squares (but not at the center of the squares). On each trial, the participants had unlimited time to make their response.

The presentation of standard and intersection positions was blocked, and the order of presentation counterbalanced over participants. Within each block (standard, intersection), the 20 random positions (4 position types \times 5 stimuli) were first presented in a different random order for each participant. The 5 game positions were then presented, also in a different random order for each participant. The participants had two practice trials for random positions and one practice trial for game positions.

Data Reduction and Analysis

To test differences in recall for the standard and intersection positions, we used a two-way layout (standard vs. intersection) \times position type (five levels: game, 1/3, 2/3, random, or 3/3) ANOVA. For the intersection positions, to obtain estimates of the skill slopes, we performed regressions in which recall performance (the dependent variable) was predicted from BCF rating, age, and VM, entered together (the independent variables). (Results of regression analyses of recall for the standard positions are reported in Gobet & Waters, 2003). As an index of chess skill, we used the participants' ratings in the BCF rating list at the time the data were collected. Following Gobet and Waters, we also included age and VM in all the models. Age has been shown to be an important variable in chess memory (Charness, 1981a, 1981b), and VM has been associated with recall performance on this task (Waters, Gobet, & Leyden, 2002).

Separate regressions were carried out for each position type. The unstandardized parameter estimate for BCF rating provided our estimate of the skill slope, and the p value of this statistic determined whether it was significantly different from zero. (To facilitate direct comparisons between the coefficients, we also will report the standardized parameter estimates.) To determine whether the skill slope for one position type was steeper than that for another position type, we tested BCF rating \times position type interactions, where position type was entered as a repeated measures variable with five levels. Since the model makes directional and unambiguous predictions about recall performance in all conditions, we used one-tailed tests

for all analyses of skill effects on recall; two-tailed tests were used for effects of age and VM on recall and for all other analyses.

To test differences in recall of bishops and knights for the intersection positions, we used a piece (bishop vs. knight) \times position type (five levels: game, 1/3, 2/3, random, and 3/3) ANOVA on percentage of recall of these two pieces. We examined the effects of skill, age, and BCF rating by adding these variables as independent variables to a regression model (piece and position type were entered as repeated measures variables with two and five levels, respectively).

Results

Percentage of Correct Recall

Table 1 shows the mean recall by skill level and position type, for both standard (left side) and intersection positions (right side). Across all skill levels, recall was worse for the intersection positions. A two-way layout (standard vs. intersection) by position type (five levels) ANOVA yielded large main effects of layout [$F(1,31) = 82.5, p < .0001$] and position type [$F(4,140) = 193.9, p < .0001$]. There was also a layout \times position type interaction [$F(4,140) = 38.8, p < .0001$], indicating that deterioration in recall for the intersection positions was most pronounced for the game positions (Table 1). However, follow-up analyses revealed that the effect of layout was significant for all the position types (all $ps < .01$).

Table 2 reports the results of regression analyses of individual position types (percentage of correct recall, intersection only). BCF rating significantly predicted recall only for the game positions; for these positions, an increase of 100 BCF grading points (e.g., the difference between an average club player and a grandmaster) yielded an increase of about 10% in recall, which corresponds to about two and a half pieces. VM significantly predicted recall for three of the nongame positions. Age tended to be negatively associated with recall, but only significantly so for the 1/3 randomized positions.

Using general linear modeling (proc glm in SAS), we also conducted an omnibus regression analysis in which all five position types were included as a repeated measures variable with five levels. This analysis revealed a significant effect of VM [$F(1,32) = 6.65, p < .05$] but no VM \times position type interaction [$F(4,128) = 0.24, p > .90$], indicating that the association between VM and recall did not differ across position types. There was also a significant effect of BCF rating [$F(1,32) = 3.05, p < .05$, one-tailed] and a significant BCF rating \times position type interaction [$F(4,128) = 2.13, p < .05$, one-tailed], indicating that the association between BCF rating and recall was significantly moderated by position type. As is illustrated in Table 2, BCF rating predicted recall for the game positions, but not for the randomized positions. There was a trend toward a significant effect of age [$F(1,32) = 3.55, p = .07$] and no age \times position type interaction [$F(4,128) = 0.57, p > .60$]. As was expected, there was also a main effect of position type [$F(4,128) = 97.9, p < .001$].

Bishop Versus Knight Recall

Table 3 shows the mean percentage of recall of bishops and knights in intersection positions by skill level and po-

Table 2
Percentages of Correct Recall: Results of Multiple Regression Analyses

Position Type	R^2	BCF Rating		Age		VM	
		b/β	SE	b/β	SE	b/β	SE
Game	15%	.10*	.05	-.25	.31	.45	.34
		.32*	.18	-.15	.18	.22	.17
One-third randomized	25%	.029	.028	-.38*	.16	.30†	.18
		.17	.17	-.40*	.17	.27†	.16
Two-thirds randomized	24%	.019	.018	-.16	.10	.26*	.11
		.18	.17	-.27	.17	.37*	.16
Random	30%	.001	.015	-.15†	.09	.27**	.10
		.01	.16	-.29†	.16	.42**	.15
Truly randomized	27%	.023†	.017	-.10	.10	.30**	.11
		.22†	.16	-.17	.17	.43**	.15

Note—For each position type, a regression was performed by entering British Chess Federation (BCF) rating, age, and visual memory (VM) in the same block as independent variables. The dependent variable was percentage recall. For each regression, the R^2 , unstandardized (upper row), and standardized (lower row) parameter estimates with standard errors are shown (for each predictor). All p values for effects of BCF rating refer to one-tailed tests; other p values reflect two-tailed tests. † $p < .10$. * $p < .05$. ** $p < .01$.

sition type. A piece (bishop vs. knight) by position type (five levels) ANOVA conducted on percentage of correct recall (for the intersection positions) revealed the expected main effect of position type [$F(4,140) = 10.0, p < .0001$]. Importantly, there was also a main effect of piece [$F(1,35) = 16.9, p < .001$], indicating that, averaged over all position types, the participants recalled bishops ($M = 11.9\%$, $SE = 0.93$) significantly better than they recalled knights ($M = 8.7\%$, $SE = 0.90$) for the intersection positions. This main effect was qualified by a significant piece \times position type interaction [$F(4,140) = 14.9, p < .0001$]. Follow-up analyses indicated that the participants were better at recalling bishops (vs. knights) for the (intersection) game and 1/3 positions ($ps < .0001$), but not for the other position types (all $ps > .05$). For comparison purposes, we also conducted a piece (bishop vs. knight) \times position type (game vs. 1/3 position) ANOVA on percentage of correct recall for the standard positions. This revealed no main effect of piece [$F(1,35) = 0.03, p > .8$], indicating that the participants were not better at recalling bishops (vs. knights) with standard (nonshifted) game and 1/3 positions.

A regression analysis on recall of bishops and knights (for the intersection positions) that incorporated BCF rating, age, and VM as independent variables yielded the expected main effect of BCF rating [$F(1,32) = 6.74, p < .01$] and a BCF rating \times position type interaction [$F(4,128) = 4.22, p < .01$], thereby paralleling the results reported above (for all the pieces). There were no significant interactions involving piece (all $ps > .05$), indicating that the effects of piece noted above were not significantly moderated by BCF rating.

Errors of Commission

Errors of commission occur when pieces are placed on incorrect squares. Over all participants, the mean numbers of errors of commission for the (intersection) game, 1/3, 2/3, random, and 3/3 positions were 5.13 ($SD = 3.70$), 3.98 ($SD = 3.98$), 4.70 ($SD = 4.63$), 4.61 ($SD = 4.49$),

and 4.74 ($SD = 4.54$), respectively. There were no significant effects of BCF rating, age, or VM. An omnibus regression incorporating position type as a repeated measures variable also yielded no significant effects of BCF rating, age, or VM.

We examined errors of commission for bishops/knights on intersection positions. We reasoned that if the participants were shifting the pieces in the mind's eye along the SE–NW diagonal, they would be more likely to misplace bishops on the SE and NW intersections of the target intersection than on the SW and NE intersections of the target intersection. Averaged over all position types, the mean percentage of bishops misplaced on the SE/NW intersections of the target intersection was 1.89 ($SD = 1.97$), and the mean percentage of bishops misplaced on the SW/NE intersections of the target intersection was 1.25 ($SD = 1.90$). This difference approached significance [$F(1,35) = 3.68, p = .06$]. The mean percentage of knights misplaced on the SE/NW and SW/NE intersections of the target intersection was 1.54 ($SD = 2.05$) and 1.46 ($SD = 2.53$), respectively. This difference was not statistically significant.

Discussion

Recall was impaired for the intersection positions, as compared with the standard positions (Table 1). Consistent with CHREST's predictions, this impairment was especially pronounced for the intersection game positions (significant layout \times position type interaction). Our working assumption is that human processing is slowed down by the processes of carrying out mental transformations to recenter the pieces, which impairs the ability to access chunks/templates in the intersection game positions.

Skill effects were present for the intersection game positions, but not for the other positions (significant skill \times position type interaction). This interaction is consistent with CHREST's predictions. Skill was not associated with errors of commission, meaning that the superior recall of better players did not come at a cost of more errors of com-

Table 3
Mean Percentages of Recall (With Standard Deviations) of Bishops and Knights by Skill and Position Type

Skill	Bishops						Knights													
	Game		1/3		2/3		Random		3/3		Game		1/3		2/3		Random		3/3	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
GMs	35.70	22.90	13.70	9.22	2.38	4.30	8.50	6.82	9.75	7.62	10.70	8.63	4.11	4.06	10.30	4.87	11.10	13.40	7.32	4.22
Experts	24.30	18.60	14.90	12.00	2.78	11.10	11.10	6.80	11.00	7.96	11.50	11.30	7.08	8.55	7.29	7.26	12.70	10.70	8.86	7.50
Class A/B	21.70	14.80	13.80	8.57	6.11	6.65	10.50	8.11	7.62	9.52	6.25	13.50	7.38	7.32	6.46	7.99	16.50	13.20	9.92	6.97
Class C/D	9.52	7.50	11.30	11.50	6.35	5.00	7.65	9.11	6.24	6.36	3.57	6.10	4.11	4.06	5.26	9.60	11.10	6.59	8.34	7.45
All (N = 36)	22.90	18.30	13.70	10.10	4.32	5.16	9.76	7.46	8.88	7.97	8.33	10.80	6.01	6.69	7.25	7.46	13.10	11.20	8.75	6.60

Note—Grandmasters (GMs), *n* = 7; experts, *n* = 12; Class A/B, *n* = 10; Class C/D, *n* = 7.

mission. This means that the skill effect was unlikely to be due to more extensive guessing by the better players.

The participants were better at recalling bishops than knights for the intersection positions (but not for the standard positions). We interpret this result to indicate the presence of mental imagery: The mental transformations were easier for the bishops than for the knights. In addition, the participants tended to be more likely to misplace bishops on the intersections SE and NW of the target intersection than on the intersections SW and NE of the target intersection. This pattern of results is consistent with the use of mental imagery along the SE–NW diagonal and provides some additional support for the presence of imagery.

The evidence for imagery was stronger for the game and 1/3 intersection positions than for the other positions (significant piece × position type interaction; see Table 3). Nonetheless, we do not rule out the possibility that some imagery was involved for all the positions. Note that mean recall of bishops/knights was poor (*M* = 8.7%) for the three most randomized positions (see Table 3). Speculatively, for the more random positions, so few bishops/knights get shifted and actually get encoded in STM (as a piece or as part of a chunk) that the bishop versus knight effect may be less visible for those positions.

HUMAN STUDY 2

Our working assumption is that the results obtained in Human Study 1 reflected imagery processes. For example, we assume that the poorer recall for the intersection (vs. standard) positions reflected the costs incurred by shifting the pieces in the mind’s eye before recognition could occur (for the intersection positions). The superior recall of bishops versus knights provided more direct evidence that imagery played a role. Nonetheless, it is also possible that the results in Human Study 1 reflected lower level processes. For example, the unfamiliar intersection positions may have required more or longer eye fixations than did the standard positions. The effect of layout (in the human data) may have, therefore, reflected the difference in how much information the participants were able to extract during the 5-sec presentation. Thus, the results in Human Study 1 may have reflected differences in visual processing, in addition to, or instead of, mechanisms of the mind’s eye.

To address the role of lower level factors, we reexamined recall data from an experiment that had high perceptual demands. Gobet and Simon (2000) included one condition in which participants had to recall positions that were presented for only 1 sec. We assume that the participants in this condition would have experienced difficulty in perceiving the entire board. Both game and random positions were presented, using the standard layout (there were no intersection positions). If the superior recall of bishops over knights in the intersection condition was due to low-level processes, the same superiority should be found in the 1-sec condition. On the other hand, if this superiority was due to mental imagery, it should not be present in the 1-sec condition, since there is no need to recenter the pieces in the mind’s eye.

Method

Participants

Twenty participants (mean age = 32.9 years, $SD = 11.6$) drawn from Gobet and Simon (2000) completed a 1-sec recall task. Their mean Elo (1978) rating was 2,131 ($SD = 256$). Further details of the participants and experimental procedures are available in Gobet and Simon (2000).

Results

Mean recall of bishops/knights for the standard game and random positions is shown in Figure 4 (bottom panel). A repeated measures ANOVA revealed the expected main effect of position [$F(1,19) = 11.7, p < .01$], indicating that recall was better for game positions than for random positions. There was no main effect of piece [$F(1,19) = 2.32, p > .1$], indicating that the participants were not better at recalling bishops than at recalling knights. There was no piece \times position interaction [$F(1,19) = 1.46, p > .1$]. A regression analysis that incorporated Elo rating as an independent variable yielded the expected main effect of Elo rating [$F(1,18) = 22.6, p < .001$], indicating that more skilled participants exhibited generally superior recall, and an Elo rating \times position type interaction [$F(1,18) = 24.0, p < .001$], indicating that skill effects were larger for the game positions than for the random positions (i.e., the typical finding). There were no interactions between piece and Elo rating (all $ps > .1$).

Discussion

Recall in the 1-sec presentation condition exhibited the expected skill \times position type interaction. However, in contrast with the results for the 5-sec intersection position, the participants were not better at recalling bishops than at recalling knights. If anything, they were (nonsignificantly) better at recalling knights. Thus, the effect of piece is specific to the 5-sec intersection condition (Figure 4, middle panel) and does not appear for either the 5-sec standard positions (Figure 4, upper panel) or the 1-sec standard positions (Figure 4, lower panel). This demonstrates that there are qualitatively different patterns of recall in the two conditions that are perceptually challenging (5-sec intersection and 1-sec standard). We argue that this qualitatively different pattern of data likely reflects the presence of imagery processes in the 5-sec intersection condition, but not in the 1-sec (or 5-sec) standard condition, at least in those conditions in which the piece effect is robust (intersection game positions).

COMPUTER SIMULATION STUDY 2

In the simulations presented above, we assumed that, during perception, pieces within a chunk had to be re-centered individually (serially) in the mind's eye before chunks could be recognized. This assumption was consistent with Kosslyn et al.'s (1988) findings, which showed that the generation of mental images is done serially. A more lenient assumption would be that a group of pieces could be shifted in parallel. We carried out simulations to test this alternative assumption. We kept the program the same as for the main simulations but assumed that a group

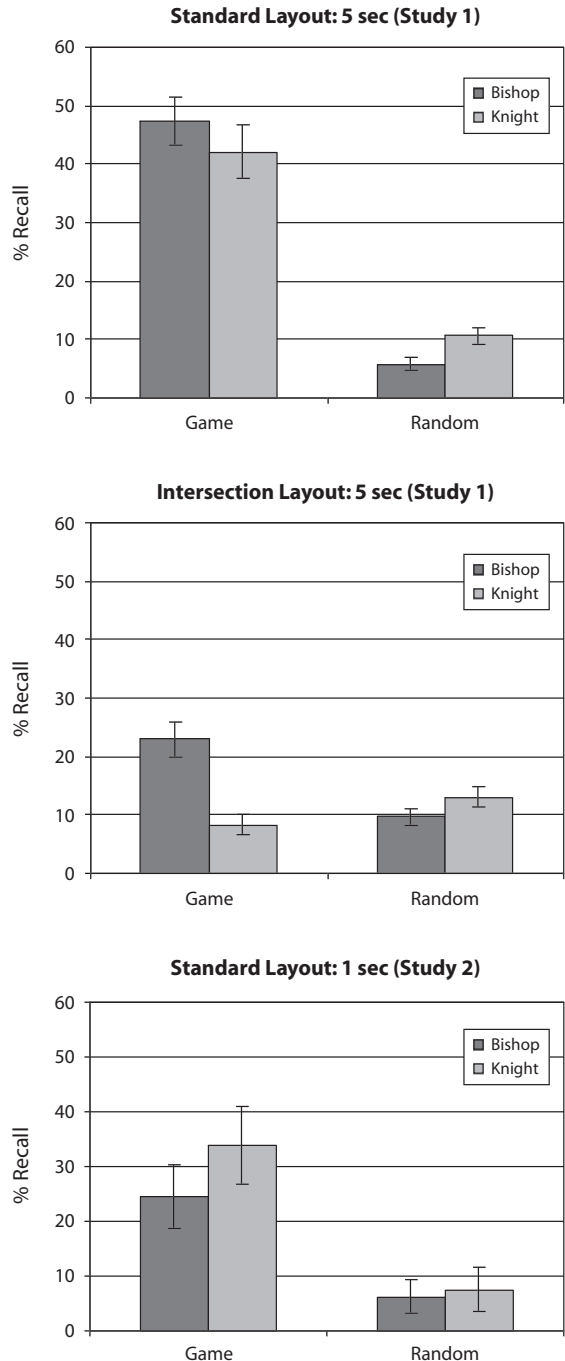


Figure 4. Recall of bishop versus knight as a function of the type of position (game or random) and type of experiment (standard layout, 5 sec; intersection layout, 5 sec; or standard layout, 1 sec).

of pieces lying within CHREST's visual field could be shifted in parallel in 125 msec.

The results of the simulation are shown in Table 1. Although the r^2 values are similar for the serial and parallel models (92% and 93%, respectively, on average), the

average absolute deviation values for the serial model were smaller than those observed with the parallel model (3.15 vs. 8.99 for the serial and parallel models, respectively). A similar outcome was observed with the *SSE* values (70.7 vs. 670.1 for the serial and parallel models, respectively). Thus, parallel shifting leads to a worse, rather than a better, fit.

GENERAL DISCUSSION

To study mental imagery and chunking, we created a new type of experimental material in which chess pieces were placed on the intersections between squares. Predictions were made by running simulations with CHREST, a computational model of expertise. We manipulated the location of pieces (standard vs. intersection), skill, and position type. The simulations assessed the effects of these manipulations on recall. An experiment with chess players assessed the accuracy of CHREST's predictions.

CHREST made several predictions that related to processes putatively carried out in the mind's eye. First, CHREST predicted that intersection positions should be recalled worse than standard positions. This effect was clearly present in the human data (significant effect of layout; see Table 1). Second, within the intersection positions, CHREST predicted that the skill difference in recall should be larger with the game positions than with the randomized positions. This was observed in the human data (significant skill \times position type interaction; see Tables 1 and 2). A skill effect was robust only for the game positions and was not observed for the randomized positions (see Table 2). Third, the time parameters of CHREST were also well supported. In general, the data supported the idea that pieces must be recentered in the mind's eye before pattern recognition can happen, and they also provided support for the assumption that the transition time is 125 msec.

Despite these successes, data from the human experiment challenged CHREST's predictions. CHREST predicted significant skill slopes for all intersection positions, but the data revealed that a significant slope was observed only for the game positions (see Table 2). It is possible that the skill effects for the randomized positions were too small to be detected reliably in our experiment (note that all the slopes had positive values with the human data, although the slope for the random positions was negligible).

The study had strengths. First, given that chess memory is a well-studied area, we used a memory task to study imagery. There was direct evidence that imagery did, indeed, play an important role in the memory task: We found that the participants were better at recalling bishops than at recalling knights for the intersection positions, presumably because they found it easier to manipulate the bishop in the mind's eye, given that the transition for the bishop (but not the knight) is congruent with its typical movement. This Stroop-like interference effect parallels that reported by Bachmann and Oit (1992), using a different imagery task, and provides additional evidence that imagery played an important role in task performance. The pat-

tern of errors of commission for bishops and knights was also supportive of imagery processes along the SE-NW diagonal line (Human Study 1).

Second, we note that CHREST makes absolute predictions about performance and that these predictions were correct in a number of cases. For example, the decrease in the recalling of game positions in the intersection condition, as compared with the standard condition, is explained by the fact that the ability to access chunks and templates, whose core consists of large chunks, is made harder. This, in turn, is explained by the assumption that carrying out mental transformations in the mind's eye to recenter the pieces has a time cost, specified by the base and square CHREST parameters.

Third, the seriality assumption was supported by the data (Simulation Study 2). We emphasize that CHREST had simulated a number of phenomena (e.g., recall of game and random positions or positions modified by mirror images) several years before the present data were collected. Although the model has several parameters, they are all set, and thus the number of degrees of freedom of the model is small. The main degree of freedom is the type and amount of input used to let CHREST acquire chunks and templates. To our knowledge, no other theory of expertise makes predictions at this level of detail.

Last, we recruited a sample of players, ranging from weak club players to grandmasters, that was of sufficient size to detect the modest associations present in the data.

The study also had a number of limitations. The imagery task we used was somewhat artificial (e.g., as compared with a check detection task). For example, in our task, the pieces needed to be mentally shifted only half a square (rather than up to seven squares), and they also needed to be mentally shifted only in the diagonal plane (and not in the horizontal/vertical plane). Strictly speaking, therefore, our results pertain only to a special case of imagery that is atypical for chess players. Thus, the generalizability of the findings is not clear. Nonetheless, it is encouraging that the timing data derived from our task converge with estimates derived from different tasks. In addition, our procedures did not allow us to address fine-grained questions about the psychological processes underlying the imagery. For example, we do not know the extent to which recentering is conducted automatically or whether it is partly under conscious control.

In addition, the simulations were conducted in a manner as close as possible to earlier simulations. Later versions of CHREST will have to (1) better capture the detail of how the mind's eye generates and maintains visual images and (2) model the differential recall of bishops and knights. Both these issues were outside the scope of the present study. Later versions may also address the import of additional assumptions (e.g., the re-placement assumption in recall).

In conclusion, the simulation and human data reported in this article have shed light on mental imagery and chunking. Perhaps the most arresting finding is that CHREST's time parameters, which were based on the sparse and somewhat inconsistent data available to de Groot and Gobet (1996), turned out to be surprisingly accurate. Al-

though previous simulations have supported CHREST's mechanisms for explaining perception and memory (e.g., de Groot & Gobet, 1996), the present article establishes the plausibility of CHREST's mechanisms for explaining mental imagery, at least in the domain of chess.

AUTHOR NOTE

Correspondence concerning this article should be addressed to F. Gobet, Centre for Cognition and Neuroimaging, School of Social Sciences, Brunel University, Uxbridge, Middlesex UB8 3PH, England (e-mail: fernand.gobet@brunel.ac.uk).

REFERENCES

- ATTNEAVE, F., & CURLEE, T. E. (1983). Locational representation in imagery: A moving spot task. *Journal of Experimental Psychology: Human Perception & Performance*, *9*, 20-30.
- AVERBACH, E., & CORIELL, A. S. (1961). Short-term memory in vision. *Bell System Technical Journal*, *40*, 309-328.
- 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.
- BADDELEY, A. (1986). *Working memory*. Oxford: Oxford University Press, Clarendon Press.
- BINET, A. (1981). *Psychologie des grands calculateurs et joueurs d'échecs* [The psychology of great calculators and chess players]. Paris: Slatkine. (Original work published 1894)
- CAMPITELLI, G., & GOBET, F. (2005). The mind's eye in blindfold chess. *European Journal of Cognitive Psychology*, *17*, 23-45.
- CHARNESS, N. (1981a). Aging and skilled problem solving. *Journal of Experimental Psychology: General*, *110*, 21-38.
- CHARNESS, N. (1981b). Search in chess: Age and skill differences. *Journal of Experimental Psychology: Human Perception & Performance*, *7*, 467-476.
- CHASE, W. G., & SIMON, H. A. (1973a). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215-281). New York: Academic Press.
- CHASE, W. G., & SIMON, H. A. (1973b). Perception in chess. *Cognitive Psychology*, *4*, 55-81.
- 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.
- DE GROOT, A. D. (1978). *Thought and choice in chess*. The Hague: Mouton. (Original work published 1946)
- DE GROOT, A. D., & GOBET, F. (1996). *Perception and memory in chess: Heuristics of the professional eye*. Assen: Van Gorcum.
- EKSTROM, R. B., FRENCH, J. W., HARMAN, H. H., & DERMAN, D. (1976). *Kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.
- ELO, A. (1978). *The rating of chessplayers, past and present*. New York: Arco.
- FERRARI, V., DIDIERJEAN, A., & MARMÈCHE, E. (2006). Dynamic perception in chess. *Quarterly Journal of Experimental Psychology*, *59*, 397-410.
- FREUDENTHAL, D., PINE, J. M., & GOBET, F. (2005). Resolving ambiguities in the extraction of syntactic categories through chunking. *Cognitive Systems Research*, *6*, 17-25.
- FREUDENTHAL, D., PINE, J. M., & GOBET, F. (2006). Modeling the development of children's use of optional infinitives in Dutch and English using MOSAIC. *Cognitive Science*, *30*, 277-310.
- GOBET, F. (1997). A pattern-recognition theory of search in expert problem solving. *Thinking & Reasoning*, *3*, 291-313.
- 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., & JACKSON, S. (2002). In search of templates. *Cognitive Systems Research*, *3*, 35-44.
- 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). Recall of random and distorted chess positions: Implications for the theory of expertise. *Memory & Cognition*, *24*, 493-503.
- GOBET, F., & SIMON, H. A. (1996b). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology*, *31*, 1-40.
- 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.
- GOBET, F., & WATERS, A. J. (2003). The role of constraints in expert memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *29*, 1082-1094.
- GRUBER, H. (1991). *Qualitative Aspekte von Expertise im Schach* [Qualitative aspects of expertise in chess]. Aachen: Feenschach.
- HOLDING, D. H. (1985). *The psychology of chess skill*. Hillsdale, NJ: Erlbaum.
- JONES, G., GOBET, F., & PINE, J. M. (2005). Modelling vocabulary acquisition: An explanation of the link between the phonological loop and long-term memory. *Journal of Artificial Intelligence & Simulation of Behaviour*, *1*, 509-522.
- KALAKOSKI, V. (2006). *Constructing skilled images* (Research Rep. 35). Helsinki: University of Helsinki.
- KOSSLYN, S. M. (1994). *Image and brain: The resolution of the imagery debate*. Cambridge, MA: MIT Press, Bradford Books.
- KOSSLYN, S. M., CAVE, C. B., PROVOST, D. A., & VON GIERKE, S. M. (1988). Sequential processes in image generation. *Cognitive Psychology*, *20*, 319-343.
- LANE, P. C. R., CHENG, P. C.-H., & GOBET, F. (2000). CHREST+: Investigating how humans learn to solve problems using diagrams. *Artificial Intelligence & Simulation of Behaviour Quarterly*, *103*, 24-30.
- LARKIN, J. H., McDERMOTT, J., SIMON, D. P., & SIMON, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, *208*, 1335-1342.
- LARKIN, J. H., & SIMON, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *11*, 65-99.
- LOGIE, R. H. (1986). Visuospatial processing in working memory. *Quarterly Journal of Experimental Psychology*, *38A*, 229-247.
- MILOJKOVIC, J. D. (1982). Chess imagery in novice and master. *Journal of Mental Imagery*, *6*, 125-144.
- NEWELL, A., & SIMON, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- PAIGE, J. M., & SIMON, H. A. (1966). Cognitive processes in solving algebra word problems. In B. Kleinmuntz (Ed.), *Problem solving: Research, method, and theory* (pp. 51-119). New York: Wiley.
- RICHMAN, H. B., STASZEWSKI, K. J., & SIMON, H. A. (1995). Simulation of expert memory using EPAM IV. *Psychological Review*, *102*, 305-330.
- RUCHKIN, D. S., GRAFMAN, J., CAMERON, K., & BERNDT, R. S. (2003). Working memory retention systems: A state of activated long-term memory. *Behavioral & Brain Sciences*, *26*, 709-728.
- SAARILUOMA, P. (1994). Location coding in chess. *Quarterly Journal of Experimental Psychology*, *47A*, 607-630.
- SAARILUOMA, P. (1995). *Chess players' thinking: A cognitive psychological approach*. London: Routledge.
- SAARILUOMA, P., & KALAKOSKI, V. (1997). Skilled imagery and long-term working memory. *American Journal of Psychology*, *110*, 177-201.
- SHEPARD, R. N., & COOPER, L. A. (1982). *Mental images and their transformations*. Cambridge, MA: MIT Press.
- SIMON, H. A. (1969). *The sciences of the artificial*. Cambridge, MA: MIT Press.
- SIMON, H. A. (1978). On the forms of mental representation. In C. W. Savage (Ed.), *Perception and cognition: Issues in the foundations of psychology* (pp. 3-18). Minneapolis: University of Minnesota Press.
- SIMON, H. A., & GILMARTIN, K. J. (1973). A simulation of memory for chess positions. *Cognitive Psychology*, *5*, 29-46.
- TABACHNECK-SCHIFF, H. J. M., LEONARDO, A. M., & SIMON, H. A. (1997). CaMeRa: A computational model of multiple representations. *Cognitive Science*, *21*, 305-350.
- VICENTE, K. J., & WANG, J. H. (1998). An ecological theory of expertise effects in memory recall. *Psychological Review*, *105*, 33-57.
- WAGHORN, K. (1988). *Chess players' use of task-specific processes in a perceptual classification task*. Unpublished honors thesis, University of Waterloo.

WATERS, A. J., GOBET, F., & LEYDEN, G. (2002). Visuospatial abilities of chess players. *British Journal of Psychology*, **93**, 557-565.

NOTES

1. Pointers might be implemented by STM neurons in the prefrontal cortex, firing in synchrony with neurons in posterior areas of the brain. Capacity of STM (i.e., the number of pointers that can be held simultaneously in STM) would then be a function of the number of distinct frequencies available (e.g., Ruchkin, Grafman, Cameron, & Berndt, 2003).

2. The British Chess Federation (BCF) rating is an interval scale ranking competitive chess players, similar to the Elo (1978) rating, a more widely used rating system. Skill levels have standard names, which are used consistently in this article as follows (the approximate corresponding range in BCF points is given in parentheses): grandmaster (normally above 240), international master (225–240), master (200–225), expert (175–200), Class A players (150–175), Class B players (125–150), and so on. There is an international World Chess Feder-

ation (FIDE: Fédération Internationale des Échecs) rating (also called an Elo rating), but usually, national federations have their own rating (e.g., BCF). The formula for converting BCF into Elo is $(BCF * 5) + 1,250$. FIDE publishes rating lists of its members every 3 months and awards titles such as grandmaster.

3. There are two ways that the knight move is typically taught, one combining horizontal and vertical movements (2 squares, 1 square; 1 square, 2 squares) and the second involving a horizontal or vertical move and then a diagonal move outward (1 square, 1 square). Thus, under certain conditions, the knight could be imagined as moving diagonally. However, the diagonal movement is clearly less closely associated with the knight than with the bishop (which moves only diagonally). In addition, we suspect that adult chess players typically represent the knight's movement as a straight line between the home and the target squares (and not in terms of horizontal/vertical/diagonal shifts).

(Manuscript received December 10, 2006;
revision accepted for publication October 25, 2007.)