Abstract—The reported research extends classic findings that after briefly viewing structured, but not random, chess positions, chess masters reproduce these positions much more accurately than less-skilled players. Using a combination of the gaze-contingent window paradigm and the change blindness flicker paradigm, we documented dramatically larger visual spans for experts while processing structured, but not random, chess positions. In addition, in a check-detection task, a minimized 3 × 3 chessboard containing a King and potentially checking pieces was displayed. In this task, experts made fewer fixations per trial than less-skilled players, and had a greater proportion of fixations between individual pieces, rather than on pieces. Our results provide strong evidence for a perceptual encoding advantage for experts attributable to chess experience, rather than to a general perceptual or memory superiority.

Simon and Chase (1973) proposed that much as Drosophila can be used as a model organism for the study of genetics, chess offers cognitive scientists an ideal task environment for the study of skilled performance. Since 1946, when de Groot (1978) conducted his pioneering investigation showing that perception and memory are more important differentiators of expertise than is the ability to think ahead in the search for good moves, chess research has been instrumental in enhancing understanding of human expertise (Ericsson & Charness, 1994) and in contributing to the study of artificial intelligence (Charness, 1992). In a classic study, Chase and Simon (1973a, 1973b) replicated and extended de Groot’s findings that after viewing chess positions for only a few seconds, chess masters were able to reproduce these positions much more accurately than less-skilled players. There was little difference in performance as a function of expertise when random board configurations were used, indicating that the superior immediate memory performance of the skilled players was not attributable to a general superiority or unique structure of their memory systems or processes (e.g., photographic memory; see Binet, 1894). Rather, Chase and Simon postulated that experts use chess knowledge to create meaningful chunks consisting of several chess pieces, and are thus able to encode structured, but not random, chess configurations more quickly and accurately. More recently, a very small but reliable advantage in recall for random configurations has been shown for more expert players, though this is probably attributable to the occasional presence of familiar chunks in random positions (Gobet & Simon, 1996a). Further illustrating the critical importance of knowledge structures for performance, Chi’s (1978) work comparing children who were skilled chess players with novice adults showed an advantage for children in chess recall, but an advantage for adults in digit recall.

Chase and Simon (1973a, 1973b) hypothesized that much of the skilled chess player’s advantage lies in the early perceptual organization and internal representation of the chess position. Specifically, they argued that the link between skilled perception and skilled problem solving was to be found in the associations between perceptual chunks and generation of plausible moves. The size of an expert’s vocabulary of chess-related configurations was initially estimated to be 50,000 to 100,000 chunks (Simon & Gilmartin, 1973), although small perceptual chunks are most likely supplemented by larger structures termed templates (Gobet & Simon, 1996b).

The master is thought to use recognizable configurations of pieces, chunks, and templates as indices to long-term memory structures that, in association with a problem-solving context, trigger the generation of plausible moves for use by a search mechanism. Search is thereby constrained to the more promising branches in the space of possible moves from a given chess position. Hence, grandmasters, the best human players, can find excellent moves despite generating only a small number of potential states (perhaps 100 or so for a few minutes of search; Charness, 1981). Such constrained search differs sharply from the enormous space explored by computer chess programs, which typically explore millions to 100s of millions of alternatives in the same time frame.

The present research employed eye movement–monitoring techniques in order to provide direct evidence for the hypothesis that a perceptual advantage is a fundamental component of chess skill. We predicted that the perceptual advantage accruing to expert chess players would be reflected in a larger visual span for chess-related visual patterns, but not for patterns unrelated to chess. Specifically, we predicted that chess experts’ encoding of chunks rather than individual pieces would result in fewer fixations, and fixations between rather than on individual pieces. Such a visual-span advantage would also mean that while examining structured, but not random, chess configurations, experts would make greater use of parfoveal processing to extract information from a larger portion of a chessboard during an eye fixation (hence the term visual span, also referred to in the literature as the perceptual span or the span of effective vision; see Jacobs, 1986; Rayner, 1998). Prior research on eye fixations in chess has also shown differences in variables such as fixation duration and coverage of the chessboard (de Groot & Gobet, 1996, chap. 6).

In the current study, we used a gaze-contingent window technique (e.g., McConkie & Rayner, 1975; see Rayner, 1998, for a review) to measure visual span as a function of chess skill (expert vs. intermediate vs. novice) and configuration type (chess configuration vs. random configuration). As shown in Figure 1, a gaze-contingent window requires obscuring the identity of all chess pieces except those within a certain “window” that is continually centered on the participant’s current gaze position. The participant’s visual span is measured by varying the size of the window over successive trials and determining the smallest possible window that does not significantly interfere with the participant’s task performance.

We combined the gaze-contingent measurement of span size with the flicker paradigm introduced by Rensink, O’Regan, and Clark.
Chessboards containing structured or random configurations were modified by changing the identity but not the color of a single piece (see Fig. 1). In each trial, images of the original and modified boards were displayed sequentially and alternated repeatedly, with a blank interval between displays. Participants had to detect the changing piece. Previous research indicated that participants are surprisingly poor at change detection in the flicker paradigm, a phenomenon termed change blindness (Rensink et al., 1997; see Simons & Levin, 1997, for a review). Note that change detection in the present task required no chess knowledge, and consequently we were able to explore visual span across a broad range of chess skill stretching from novice to master. We predicted that when processing chess configurations, but not random configurations, chess experts would demonstrate larger visual spans and better change detection than less-skilled players.

To examine differences in the spatial distribution of fixations be-
tween experts and novices, we monitored eye movements of another sample of chess players in a check-detection task. Saariluoma (1985) showed that master players can rapidly and accurately decide whether a chess piece is attacked, and do so more quickly than their less-skilled counterparts. The rather simple chess relation of check detection (attack of a King) is highly salient and presents a good model for the extraction of chess-relevant relations among pieces. As shown in Figure 2, in the present study, check detection was performed using a minimized 3 × 3 chessboard containing a Black King and one or two potentially checking pieces. At the beginning of each trial, participants fixated the center square of the board, a square that was always empty. A large visual span in this task may result in few if any saccades (the rapid eye movements that shift the point of gaze and define the beginning and end of fixations) during a trial and in fixations between, rather than on, individual pieces. To demonstrate that the encoding advantage of experts is related at least in part to their chess experience, rather than to a general perceptual superiority, we manipulated the familiarity of the notation (symbol vs. letter) used to represent the chess pieces. The symbol and letter notations were used to represent identical chess problems. However, the symbol representation is much more familiar than the letter representation. Consequently, if encoding efficiency is related to chess experience, any skill advantage would be more pronounced in the symbol than in the letter trials (i.e., a skill-by-notation interaction).

METHOD

Visual Span in the Flicker Paradigm

Participants

Thirty-two paid participants (16 novices, 8 intermediate players, and 8 experts) were included in this task. All participants had normal or corrected-to-normal vision. Chess Federation of Canada (CFC) ratings for the expert players ranged from 2200 to 2400 ($M = 2278$). CFC ratings for the intermediate players ranged from 1300 to 1700 ($M = 1483$). The mean rating in the CFC is about 1600, with a

Fig. 2. Illustration of the stimuli used in the check-detection task, with example scan paths superimposed (numbers represent duration in milliseconds). An example is shown for each skill group and notation condition.
standard deviation of about 200. Players ranged in age between 18 and 33 years. The novices were inexperienced chess players who typically reported playing no games of chess in the past year and very few games over their lifetime. Informed consent was obtained, and the rights of the participants were protected.

**Materials**

The stimulus displays showed chessboards subtending a visual angle of 12.8° horizontally and vertically, and including chess pieces approximately 1.3° in diameter. Two types of configurations were used: chess configurations (with 20 chess pieces in each) selected from a large database of chess games and random configurations, which were created by repeatedly and randomly exchanging pieces in the chess configurations. Thus, each random configuration maintained the spatial configuration of the chess position from which it was derived, but destroyed the chess relation information. Each configuration was presented in both its original form and a modified form in which the identity but not the color of a single piece was changed (see Fig. 1). In every trial, the original and modified versions were presented in alternation, for 1,000 ms each, with a 100-ms blank screen presented between displays. This sequence continued until a decision was made. The participant’s task was to detect the changing piece. In the trials using a gaze-contingent window, the pieces outside a circular, gaze-centered window were replaced with gray blobs masking the actual colors and shapes.

**Apparatus**

Eye movements were measured with an SR Research Ltd. EyeLink system. Following calibration, gaze-position error was less than 0.5°. The temporal resolution of the system was 4 ms. In order to minimize the delay between physical eye movements and updates of the display (average delay = 14 ms), we monitored eye movements exclusively for the purpose of controlling the gaze-contingent window; they were not otherwise recorded or analyzed.

**Procedure**

Prior to every trial, participants were asked to fixate a marker in the center of the display. Following a button press, an experimental display was presented. As soon as participants detected the changing piece (the target), they ended the trial by pressing another button and naming the alternating pieces. The experimenter monitored and recorded the accuracy of their performance. There were very few errors in this task across all groups and conditions (error rate < 1%). After 8 practice trials for each of the two conditions (chess configurations vs. random configurations), the experiment started with 16 baseline trials for each condition. In these trials, no gaze-contingent window was used (i.e., all pieces were visible throughout the trial). The normative reaction time (RT) for each condition was calculated as the third quartile of the baseline RT distribution.

Following the baseline trials, 24 blocks with eight RT measurements in each block (targets appearing twice in each of the board’s quadrants) were presented. Each block consisted of trials in one of the two experimental conditions (chess vs. random configurations), and conditions were alternated across blocks. The RT in each block was used to determine the window size of the next block in the same condition. For the first block in each condition, window size was set to 8° in diameter (representing approximately 19 squares). The window was centered on participants’ gaze position and moved following a change in gaze position. The window size was increased if the median of the RT within a block was longer than 102% of the normative RT and decreased if it was shorter than 98% of the normative RT. The first adjustment in each condition was an increase or decrease of 1.28°, and each successive adjustment was 9% smaller than the one preceding it. Consequently, the final adjustment in the sequence of 12 blocks in each condition represented a change of only 0.45°. The visual span for each condition was calculated as the mean of the last two window sizes (i.e., the window sizes after the 11th and 12th adjustments).

To account for any changes in performance over time due to practice or fatigue, the normative speed was updated after every sequence of three adjustments. This was accomplished by administering eight baseline trials for each condition. These trials replaced the eight earliest baseline trials, and the normative speed was recomputed.

**Check Detection**

**Participants**

Forty paid participants (20 novices, 10 intermediate players, and 10 experts) who did not perform the change-detection task in the flicker paradigm were included in this task. All participants had normal or corrected-to-normal vision. CFC ratings for the expert players ranged from 1950 to 2352 (M = 2117). CFC ratings for the intermediate players ranged from 962 to 1387 (M = 1226). Players ranged in age between 19 and 31. The novices reported playing no games of chess in the past year and very few games over their lifetime. Informed consent was obtained, and the rights of the participants were protected.

**Materials**

A minimized 3 × 3 chessboard subtending a visual angle of 9° horizontally and vertically was displayed; chess pieces on the board were approximately 1.8° in diameter. Each display contained a Black King in the top left or top right square and one or two potentially checking pieces (from the combinations of Rook, Queen, and Knight). There were no cases of a double check with two attackers, and the Queen never delivered a check on the diagonal. The center square was never occupied. The symbol notation was composed of chess symbols similar to those used for chess diagrams in chess books and magazines. Letter notation was shown in a bold sans serif font using all capitals: Q (Queen), R (Rook), N (Knight), and K (King); the Black King was a filled letter, and the other pieces were represented by outline letters (similar to the white chess pieces; see Fig. 2). Check status, spatial layout (i.e., the positioning of the King and the number and position of the attackers), and notation were completely crossed, but only the latter was analyzed. The counterbalancing and data reduction were done in order to provide sufficient power for the analysis of the spatial distribution of fixations, which was the focus of the present investigation.

**Apparatus**

The same apparatus was used as in the flicker paradigm.
Procedure
Prior to every trial, participants were asked to fixate a marker in the center of the display. Following a button press, an experimental display was presented. A trial was terminated as soon as the participant made a yes/no response regarding the check status by pressing one of two response buttons. After 64 practice trials during which participants could ask any questions about the task or symbols, 384 experimental trials were administered (192 in each of the notation conditions).

RESULTS

Visual Span in the Flicker Paradigm

Figure 3a displays the average median RTs obtained in the non-gaze-contingent baseline trials that were used to compute the normative RT criteria. As the figure indicates, the difference between RTs for chess versus random configurations was significant only in the expert group: experts—\( t(7) = 5.4, p < .001 \); intermediate players—\( t < 1 \); novices—\( t < 1 \). Thus, there was a significant skill-by-configuration-type interaction, \( F(2, 29) = 9.93, MSE = 3.0, p < .001 \). Furthermore, for random-configuration trials, RTs did not differ significantly across the skill groups, \( F(2, 29) = 6.93, MSE = 2.268,085, p < .01 \), with experts being significantly faster than both intermediate players and novices.

The visual-span results shown in Figure 3b follow the same pattern as the RT results. The skill-by-configuration-type interaction was significant, \( F(2, 29) = 9.64, MSE = 38.9, p < .001 \), with experts’ span area for chess configurations being dramatically larger than the span areas in all other skill-group-by-configuration-type cells (all \( ts > 3.45, p < .01 \)), which in turn did not differ from each other (all \( ts < 1 \)).

Consistent with Chase and Simon’s (1973a, 1973b) hypothesis, the increases in visual span and speed of responding that characterize expert performance on trials with chess, but not random, configurations clearly indicate an encoding advantage attributable to chess experience, rather than to a general perceptual or memory superiority. The results are especially impressive considering that the spatial layout was identical for chess and random configurations, with only the identity of pieces being different (i.e., pieces in chess configurations were randomly exchanged to derive the random configurations).

Check Detection

The error rate was less than 5% for all combinations of skill and notation. Experts made fewer errors (1.3%) than intermediate players (2.6%) and novices (3.7%), \( F(2, 37) = 3.36, MSE = 11.9, p < .05 \) (for expert vs. intermediate and expert vs. novice, \( ts > 2.34, ps < .05 \)). There were also fewer errors for the symbol (1.9%) than the letter (3.7%) notation, \( F(1, 37) = 16.04, MSE = 2.7, p < .001 \). The skill-by-notation interaction was not significant, \( F(2, 37) = 1.21, MSE = 2.68, p = .31 \).

Check-detection RTs demonstrated a pattern similar to the one for error rates. RTs were faster for experts (861 ms) than for intermediate players (1,087 ms) and novices (1,207 ms), \( F(2, 37) = 8.51, MSE = 94,165, p < .001 \) (for expert vs. intermediate and expert vs. novice, \( ts > 2.42, ps < .05 \)). RTs were also faster for the symbol (1,036 ms) than the letter (1,145 ms) notation, \( F(1, 37) = 89.66, MSE = 2,643, p < .001 \). The skill-by-notation interaction was not significant, \( F(2, 37) = 2.58, MSE = 2,643, p = .09 \). The effect of notation on check-detection performance validates its effectiveness as a manipulation of the familiarity of the representation of chess configurations. In addi-

Fig. 3. Results for the flicker paradigm. Median reaction times in the non-gaze-contingent baseline trials (a) and area of visual span (number of squares) (b) are shown separately for the three skill groups and two configuration types. Error bars indicate standard errors.
tion, skill clearly influenced check-detection performance, validating the relevance of this task to the study of expertise in chess.

The scattergrams in Figure 4 show the spatial distributions of gaze positions in the check-detection task for the novice, intermediate, and expert groups. Given that the number of fixations varied substantially across skill groups, we scaled the size of the dots in each scattergram to make them proportional to the number of fixations in that group. An inspection of the scattergrams collapsing across all trial types (i.e., the spatial layout of chess pieces, check status, and notation), with initial gaze positions included (the top row of Fig. 4), reveals a greater concentration of dots in the center of the scattergram for the experts than in the scattergrams for the intermediate players and novices. This center-of-gravity effect reflects a large disparity between skill groups in the percentage of trials without an eye movement (i.e., without a saccade; see the example for an expert with a display using symbol notation in Fig. 2). In such trials, the gaze position remained in the center square of the chessboard throughout the duration of the trial. The percentage of trials without an eye movement was 15.9 for experts, 2.6 for intermediate players, and 1.6 for novices, $F(2, 37) = 4.58, MSE = 0.84, p < .01$ (for expert vs. intermediate and expert vs. novices).

![Fig. 4. Scattergrams of gaze positions in the check-detection task by skill. The top row presents data collapsed across all trial types and spatial layouts. The middle row presents the same data as the top row, excluding initial gaze position. The bottom row presents data collapsed across trials in which the King and a single attacker were positioned in opposite corners of the same column, excluding initial gaze position. Note that each dot represents an individual gaze position, and the size of the dots is scaled in each scattergram to make them proportional to the number of fixations in that group. A = position of an attacker piece; K = position of the King.](image)
novice, \( t > 2.14, p < .05 \). On trials in which eye movements occurred, average fixation duration did not differ as a function of skill (274 ms for experts, 269 ms for intermediate players, and 281 ms for novices), \( F(2, 37) < 1 \). On such trials, experts tended to make shorter saccades than less-skilled players (2.98° for experts, 3.43° for intermediate players, and 3.62° for novices), \( F(2, 37) = 5.04, MSE = 0.27, p < .05 \) (for expert vs. intermediate and expert vs. novice, \( t > 2.10, p < .05 \)).

Furthermore, as shown in Figure 5a, on trials in which eye movements occurred, experts made fewer fixations than intermediate players and novices, \( F(2, 37) = 8.84, MSE = 0.85, p < .001 \). More important, the skill-by-notation interaction was significant, \( F(2, 37) = 13.30, MSE = 0.02, p < .001 \); the symbol notation resulted in fewer fixations than the letter notation for both experts, \( t(9) = 6.10, p < .001 \), and intermediate players, \( t(9) = 4.53, p < .001 \), but not for novices, \( t < 1 \).

In order to compare the spatial distribution of fixation positions across groups, we computed the proportion of fixations landing on squares containing chess pieces (henceforth referred to as proportion on pieces). The scattergrams with initial gaze positions excluded (the middle and bottom rows of Fig. 4) clearly indicate that experts made proportionately fewer fixations on pieces than did intermediate players and novices, \( F(2, 37) = 5.76, MSE = 0.04, p < .01 \). As shown in Figure 5b, the skill-by-notation interaction was significant, \( F(2, 37) = 13.30, MSE = 0.02, p < .001 \); the symbol notation resulted in fewer fixations on pieces than did the letter notation for both experts, \( t(9) = 3.80, p < .01 \), and intermediate players, \( t(9) = 5.34, p < .001 \), but the opposite was the case for novices, \( t(19) = 2.61, p < .05 \). Thus, consistent with Chase and Simon’s (1973a, 1973b) chunking hypothesis, in the check-detection task, chess experts made fewer fixations than less-skilled players and placed a greater proportion of fixations between individual pieces, rather than on pieces. The magnitude of these effects was stronger for the more familiar symbol notation than for the letter notation, demonstrating that the experts’ encoding advantage is related at least in part to their chess experience, rather than to a general perceptual superiority.

**DISCUSSION**

De Groot (1978) and Chase and Simon (1973a, 1973b) showed that the chess master has an advantage in immediate memory for chess-related information following a very brief exposure to an unfamiliar position. Our study extends these findings by showing that experts have an advantage in extracting perceptual information in an individual fixation. For check detection, a task that is well defined and for which positional uncertainty is minimized, the expert extracts the necessary interpiece relations from both foveal and parafoveal regions. The larger visual span of experts in this task results in fewer fixations per trial, and a greater proportion of fixations between, rather than on, individual pieces.

The combination of the gaze-contingent window and the flicker paradigms introduced in the present study allowed for a more direct and conclusive demonstration of perceptual superiority as a function of expertise in chess. Specifically, advanced chess skill attenuates change blindness by improving target detection in meaningful, but not scrambled, chess configurations, and this effect is due to greater span size (relative to less-skilled players) in the former, but not the latter, condition (for other findings of semantic effects on change blindness, see Hollingworth & Henderson, in press; Rensink et al., 1997; Werner & Thies, in press).

Furthermore, the manipulation of notation in the check-detection task, which kept the semantics constant while changing the surface representation of a chess problem, and the manipulation of configuration type (i.e., chess vs. random) in the flicker paradigm provided powerful demonstrations of the effects of familiarity on perception. As has been found with other visual context effects (e.g., word, letter, object, face, and scene superiority effects; see Reingold & Jolicoeur,
1993), a coherent and familiar context (i.e., a chess configuration) enhances the perception of constituent elements (i.e., the identity of pieces and interpiece relations).

Finally, given the pivotal role played by eye movement paradigms in the study of reading skill (see Rayner, 1998, for a review), it is surprising that very few empirical studies have employed these techniques in chess (de Groot & Gobet, 1996; Ellis, 1973; Jongman, 1968; Tikhomirov & Poznyanskaya, 1966; Winikoff, 1967). The present study illustrates that eye movement paradigms may prove invaluable in supplementing traditional measures of performance such as RT, accuracy, and verbal reports as a means for understanding human expertise in general and chess skill in particular.

Acknowledgments—Preparation of this article was supported by a grant from the Natural Sciences and Engineering Research Council of Canada to Eyal Reingold and by National Institute on Aging Grant 5R01 AG13969 to Neil Charness.

REFERENCES


Rensink, R.A., O’Regan, J.K., & Clark, J.J. (1997). To see or not to see: The need for attention to perceive changes in scenes. Psychological Science, 8, 368–373.


(Eceived 3/31/00; Accepted 5/22/00)