Ref F #381 10

Perception in Chess

William G. Chase and Herbert A. Simon

What does an experienced chess player "see" when he looks at a chess position? By analyzing an expert player's eye movements, it has been shown that, among other things, he is looking at how pieces attack and defend each other. But we know that he is seeing much more. Our work is concerned with just what the expert chess player perceives.

It is known that chess masters search through about the same number of possibilities as weaker players, but they are very good at coming up with the "right" moves for further consideration, whereas weaker players spend considerable time analyzing the consequences of bad moves.

An intriguing difference, however, appears to exist between masters and weaker players in short-term memory. Masters show a remarkable ability to reconstruct a chess position almost perfectly after viewing it for only 5 seconds, although their short-term memory for random placement of chess pieces is no better than that of weaker players. Their superior performance with "meaningful" positions must lie in their ability to perceive structure in such positions and encode them in chunks. Specifically, if a chess master can remember the location of 20 or more pieces on the board, but has space for only about five chunks in short-term memory, then each chunk must be composed of four or five pieces, organized in a single relational structure. Our concern here is to discover and characterize the structures, or chunks, that are seen on the board and stored in short-term memory.

Previous studies of chess perception provide no direct methods for delimiting the chunk boundaries or detecting tne relations that hold among the components of a chunk. Evidence is needed on these points in order to discover how many pieces typically constitute a chunk, what the relative sizes are of the chunks of masters and weaker players, and how many chunks players retain after a brief view of a position.

The central objective of this study, then, is to isolate and define the chunks into which information is hypothesized to be encoded in chess perception tasks. We use two techniques, a *perception task* and a *memory task*. By using two different tasks, we obtain some protection against artifacts that might compromise the interpretation of our findings. One important question we shall investigate is whether the chunks defined by the data from the perception task are essentially of the same size and character as the chunks defined by the data from the memory task.

Chess Players' Tasks

Three chess players—a master (M), a Class A player (A), and a beginner (B)—were used as subjects. Twenty games between advanced players were selected from chess books and magazines to generate the stimuli. Ten were middle game positions, at about White's 21st move, with 24–26 pieces remaining on the board. Ten were end-game positions, at about the 41st move, with 12–15 pieces remaining on the board. Not all the positions were "quiet," i.e., some of them caught games at a point where an exchange of pieces was in progress.

In addition to these positions from actual games, eight random positions were generated, four from middle games and four from end games, by taking actual positions and replacing the pieces randomly on the board.

Perception task—In this task, two chess boards were placed side by side, separated by about 6 inches: One of the 28 chess positions was set up on the subject's left, and the other board, free of pieces, was placed directly in front of him. A full set of pieces was placed to the right of the blank board. A partition between the two boards prevented the subject from seeing the position on the left. When the partition was removed, the subject's task was to reconstruct the position on the board in front of him as quickly and accurately as possible, glancing at the position on the left as often as he wished. His behavior was recorded on videotape.

Memory task—The boards were set up exactly as in the perceptual task. When the partition was removed, the subject was allowed to view the position on the left for 5 seconds, and the partition was then placed in position again. The subject then recalled, by placing pieces on the board in front of him, what he could remember of the position on the left, being allowed as much time as he wished (subjects rarely took more than 1 minute). If the position was not reconstructed perfectly, the board in front of the subject was cleared and a second trial was conducted in the same way: 5 seconds of viewing, followed by free recall of the position. Additional trials followed until the subject recalled the position perfectly, except for the random positions, which were too difficult to continue to criterion.

In both the perception and memory task, each subject processed five middle-game positions, five end-game positions, two randomized middle-game positions, and two randomized end-game positions.

Analysis of Performance

Videotape records for both tasks were analyzed by recording each piece as it was placed on the board, and by recording the time, within 1/10 second, between the placing of that piece and the next one.

The time intervals were used to segment the protocols in order to test the hypothesis that long pauses would correspond to boundaries between successive chunks, while short time intervals between pieces would indicate that the pieces belonged to the same chunk in memory.

The nature of the chess relations between successive pieces, separated by long and brief pauses, respectively, were analyzed for information that would reveal how pieces are chunked perceptually. The occurrence of each of five chess relations between successively placed pieces was recorded: *attack*—either one of the two pieces attacks the other; *defense*—either one of the two pieces defends the other; *proximity*—each piece stands on one of the eight squares adjacent to the other; *common color*—both pieces are of the same color; and *common type*—both pieces are of the same type.

Accuracy of reconstruction—The accuracy with which the subjects reconstructed positions on the first trial in the memory task was analyzed for comparison with previous research findings. Accuracy was measured by the number of pieces placed on the correct squares of the board on the first trial after a 5-second view of the board. The number of pieces correct on subsequent trials was also computed, but chief interest for our purposes centers in the first-trial results.

In the actual middle game positions, M was able to place an average of about 16 pieces correctly on the first trial, while A and B placed about eight and four, respectively. M reproduced the board perfectly in three or four trials, while A typically required about one or two more trials than M, but B took considerably more trials. M showed no such superiority in additional pieces placed in successive trials. In trials just beyond the first, M typically added about four more pieces to his previous reconstruction, while the gains for A and B averaged five or six pieces per trial. This difference disappears, however, when the learning curve reaches the level of M's first-trial performance.

In the end-game positions, M placed an average of about eight pieces correctly on trial 1, while A and B placed about seven and four, respectively. In these positions, M required two or three trials to reconstruct the positions perfectly; A, about three or four; and B, between four and seven trials. Thus, in both middle- and end-game positions from actual games, ability to retain information from a 5-second view of the board was closely related to playing strength.

In the random unstructured positions there was no relation at all between memory of the position and playing strength. Moreover, the first-trial performances of all three subjects on the random positions was even poorer than B's performance on the actual game positions.

On the hypothesis that memory of positions depends on recognizing familiar configurations or chunks of pieces, a grandmaster or master would find it easier to remember positions like those he encounters in his play and study. Our subject, M, when interviewed after the experiment, reported that he was troubled by positions that looked "unreasonable." He also reported difficulty with positions that were not quiet, complaining that he couldn't get the "sense" of the position when it was in the middle of an exchange.

Accordingly, our subjects were tested on nine new positions taken from a book of chess puzzles from actual master games. Although the positions were tactical in nature, they were not in the middle of an exchange. For each subject, nine positions were chosen at random, and a single 5-second trial was conducted. For these new positions, B, A, and M averaged 33, 49, and 81 percent correct, respectively, as compared to 18, 34, and 62 percent, respectively, on the first trial of the previous positions. These data unequivocally replicate previous findings in the area of chess perception.

One unexpected result deserves note at this point. M recognized four of the nine new positions, and always within the first second of exposure, yet M's performance was virtually identical for recognized versus unrecognized positions: 83 vs. 79 percent, respectively. Also, for one of the previous middle-game positions, M suddenly recognized the game after he had placed the pieces on trial 1. This discovery did not, however, improve his recall of the position in any way.

Time intervals—In the perception task, the first thing to look at is the distribution of times between successive pieces placed on the board. These times were analyzed separately for: *within-glance intervals*—intervals between pieces placed without looking back at the original position; and *between-glance intervals*—intervals between two pieces separated by a glance back at the original position.

The results are straightforward and roughly the same, with one exception, for all three subjects. Within-glance intervals seldom exceeded 2 seconds, and the modal intervals were ½ second or less. For the between-glance intervals, there was a significant tendency for the better players to take less time.

In the memory task there is no observable behavior that corresponds to the within-glance, between-glance distinction. If we wish to compare the time intervals for the two tasks, we must use the combined frequency distribution for the perception task.

The distributions of time intervals for the two tasks are not dissimilar. In the perception task, there is a preponderance of intervals under 2 seconds but a "tail" of longer intervals. In the memory task, there are numerous intervals up to about 2½ seconds, and again a tail of longer intervals. The very short intervals in the distributions, ½ second or less, are almost all cases where the subject picked up more than one piece of a kind (pawns or rooks) at once, and placed them on the board in rapid succession. In general, it took at least 1 second to retrieve a piece from the side of the board.

The similarity of the distributions encourages us to consider the following hypothesis about the nature of the perceptual chunks:

- The pieces placed on the board by the subject in the perception task after a single glance correspond to a single chunk. About 2 seconds is required to recognize a chunk and store a label for it in short-term memory. Since short-term memory appears to have a relatively fixed capacity, measured in chunks, it is most reasonable to assume that what is held in short-term memory is not the content of the chunks, but an identifier (label) that allows the content, in long-term memory, to be located and assessed. When the label of a chunk is held in short-term memory, successive elements of the chunk can be recovered from long-term memory in some hundreds of milliseconds.
- A sequence of pieces placed on the board by the subject in the memory task with intervals of less than 2 seconds between successive pieces corresponds to a single chunk. The times required for the underlying processes are essentially the same in the memory task as in the perception task.

If our hypothesis is correct (that time intervals of 2 seconds or more correspond to boundaries between chunks) then an examination of the chess relations between successive pieces within single chunks should show these relations to be quite different from the relations between successive pieces across chunk boundaries. Furthermore, if we are right in equating the significance of long and short time intervals in the two distinct tasks (perception and memory) then the within-chunk and between-chunk chess relations in the perception task should be highly similar to the corresponding relations in the memory task.

Chess relations: *Perception task*—For each subject, the within-glance probabilities and mean interpiece latencies were calculated for each of the 16 possible combinations of attack (A), defense (D), same color (C), same piece (S), and proximity (P) relations.

The data are quite similar for all subjects: Latencies show the same systematic trends, and, for the probabilities, the product movement correlations between subjects are quite high: M vs. A = .93, M vs. B = .95, and A vs. B = .92. The same is true for the between-glance data, and the correlations for the probabilities are about the same size: M vs. A = .89, M vs. B = .89, and A vs. B = .90. Thus, the same kinds and degrees of relatedness between successive pieces holds for subjects of very different skills.

Quite different patterns emerge for the within-glance and between-glance probabilities. The between-glance probabilities are much closer to the chance levels than are the within-glance probabilities. In contrast, the within-glance probabilities are higher than chance for pairs of pieces with several relations, and lower than chance for pairs with few relations. In particular, the relations AP, DC, DPC, PCS, and DPCS have high probabilities; C, S, and null (—) relations have lower-than-chance probabilities.

These probabilities are informative about the underlying structures that the subjects are perceiving. The relation DPCS is almost totally composed of pawn chains, and the relation PCS consists almost totally of rows of pawns on the same rank. These two relations also have much shorter latencies than the others. The relation DPC consists of pieces placed on adjacent squares which have a defense relation, and the relation DC consists simply of a defense relation which also implies the same-color relation. Low frequencies for the A relation suggest that attacks are noticed only if the pieces are in close spatial proximity. The C, S, and null relations are low because subjects are placing pieces which usually have multiple relations. Thus, from the within-glance relations, it appears that subjects are noticing the pawn structure, clusters of pieces of the same color, and attack and defense relations over small spatial distances.

There is some indication from the between-glance probabilities that subjects are looking back at the chess position in order to complete some partially forgotten information or to obtain new information about a partially completed structure. Subjects also report that sometimes they look back at the chess position for specific partial information. But the striking thing about these data is that between-glance frequencies are much closer to the chance level than within-glance frequencies.

For the between- and within-glance data for the randomized positions, the pattern of probabilities was still the same across subjects. Interestingly, the data look very similar to the data from real positions. Frequencies of the PCS, DPC, and AP relations are higher than chance and of the S, C, and null relations are lower than chance for the within-glance relations, whereas frequencies of the between-glance relations are very close to chance. Apparently, subjects are noticing the same kinds of structures in the random positions as in the game positions even though such structures are rare in the random positions.

The perception experiment procedure offers no guarantee that the subject did not pick up more than one chunk at a glance, but subjects reported that it was most comfortable to glance frequently at the board and not to retain much information in short-term memory.

Chess relations: *Memory task*—The memory data for individuals again show similar patterns of latencies and probabilities for all subjects, and the correlations are about the same as in the perception data: M vs. A = .91, M vs. B = .95, and A vs. B = .95.

The first question of interest concerning the memory data is the relationship between interpiece latencies and the perceptual chunks: What evidence is there that pauses are associated with retrieval of new structures?

Longer latencies are associated with fewer interpiece chess relations. The lowest latencies occur for pawn formations (PCS and DPCS) and for pairs of rooks or pairs of knights that mutually defend each other (DCS). The other relation that occurred much more than chance was that of adjacent pieces that have a defense relation (DPC), although these latencies are relatively long. It seems clear, however, that if there is a long pause in the recall, the pieces are not likely to be closely related. If the hypothesis that time intervals of 2 seconds or more correspond to boundaries between chunks is correct, the chess relations with latencies greater than 2 seconds ought to look like chance occurrences, whereas the relations occurring within 2 seconds ought to show even more structure. Examination of the memory data partitioned into relations for latencies less than (or equal to) 2 seconds, and chess relations for latencies greater than 2 seconds shows that the hypothesis is essentially correct.

For the long pauses, the only relation that is considerably above chance is that of adjacent pieces with a defense relation (DPC). Apparently, a chunk isn't retrieved from memory completely at random. Subjects use the partially constructed board to retrieve new information and the new information often consists of the DPC relation. Also it is clear from the subjects' verbal reports and from watching subjects that the overall recall pattern is systematic, e.g., counterclockwise or clockwise recall, and that local proximities are very important.

A second hypothesis is that the short and long time intervals of the memory task have the same meaning as the within- and between-glance distinctions, respectively, of the perception task. The similarity of these patterns becomes evident when we lay the probabilities side by side and contrast them with the probabilities of their occurrence by chance. There are some slight differences between the perceptual and memory probabilities, but these differences are everywhere small compared to their differences with chance probabilities. A matrix of correlations shows two clusters: First, the within-glance probabilities (perception task) from actual game positions are highly correlated with the probabilities for the short pauses in the memory task, and the within-glance probabilities from random games are moderately correlated with these two. Second, the between-glance probabilities in random positions, between-glance probabilities in game positions, probabilities for long pauses of the memory task, and chance probabilities are all highly intercorrelated.

On the basis of these data, it is reasonable to conclude that the time intervals in the two variants of the experiment, perceptual and memory, have basically the same information processing significance. The processes that occur during an interval of more than 2 seconds between the placing of two pieces appear to be significantly different from the processes that occur during an interval of less than 2 seconds. Moreover, the nature of the differences in frequencies of relations in the two cases makes it reasonable, at least tentatively, to apply the term "chunk" to the set of pieces placed on the board in either experiment within the boundaries of a pair of long time intervals.

One final comparison between the perception and memory task concerns the chunk size. It appears that the chunks are about the same size in both perceptual and memory tasks, but that chess skill is reflected in the speed with which chunks are perceived in the perception task and the size of the chunks in the memory task.

Chunk Size and Memory Span

Having segmented the recall protocol into chunks, we now can test the hypothesis that recall is limited by the number of chunks that can be held in short-term memory. We interpret this hypothesis to mean that M's superior recall should be associated with larger chunks, but that the number of chunks should be a small constant within the memory span (7 ± 2) for all subjects.

One problem with this analysis must

be dealt with first: The recall protocols generally consist of two phases: an initial recall phase, followed by a reconstruction phase. The subjects generally placed first those groups of pieces they remembered well, then searched memory for additional pieces. During the first phase chunks tended to be relatively large and errors relatively few. During the second phase pieces tended to be placed one by one, time being taken for deliberation between pieces. Errors were relatively frequent, and in many instances the player appeared to be determining where pieces ought to be rather than recalling where he had actually seen them. This behavior was more true of M than the other subjects.

To avoid inflating our estimate of the number of chunks, we need a way of distinguishing the recall phase from the reconstruction phase. To identify the reconstruction phase, we adopted the criterion of an extremely long pause (10 seconds or more) followed by mostly errors, or a series of long pauses (5 seconds or more) with errors.

We observed, first, that the chunk size is related to chess skill for the first few chunks, but that this difference disappears in later chunks of the protocol. This relation is less true of the end-game positions, and chunks are also smaller for the end games. The middle game-end game difference simply reflects the fact that end games are less structured than middle games.

The gradual drop in chunk size during recall could be due to several things: It may be that subjects simply recall their larger chunks first. Recall, having an interfering effect on short-term memory, may cause large chunks to break up into smaller chunks as some of the relations are forgotten. Later chunks may be contaminated by some of the piece-by-piece reconstructions that are missed by our criterion; perhaps the first guesses are the best and are more likely to be correct.

The average number of chunks for each subject is well within the memory span, as hypothesized; but, contrary to our expectation, the number of chunks is related to chess skill.

Taken at face value, these data suggest that M achieves his superior performance by recalling both more chunks and larger chunks. But we know by the performance on randomized positions that M does not have a superior memory capacity.

Where, then, do these extra chunks come from? There are at least two possibilities: It may be that M does not store a small number of unrelated chunks in short-term memory; he may be able to organize the chunks on the board in some as yet undetermined way so that more chunks can be stored. Another possibility is that M is reconstructing part of the position from his general knowledge of such positions, and our criterion for these reconstructions doesn't pick up all these responses because they are more likely to be correct for M than for the other players.

In summary, the data on chunk size and memory span confirm the hypotheses that chunk size is larger for more skilled chess players, and that the number of chunks is within the memory span. However, the hypothesis that the number of chunks is invariant over different levels of chess skill is not supported.

Perceptual Chunks of a Chess Master

The sizes of successive chunks for the five middle-game and nine puzzle positions for trial 1 of the memory experiments were examined for subject M. The great bulk of the 77 chunks (two or more pieces within 2 seconds) in these 14 positions belong to a very small number of types. Of the 77 chunks, only 17 couldn't be classified into the following three categories: pawn chains, castled-king positions, or clusters of pieces of the same color. These categories are not mutually exclusive since some chunks contain more than one category. The point is, however, that over 75 percent of M's chunks belong to only three types of chessboard configurations, all highly familiar and stereotyped.

One further analysis was carried out on M's protocols. From an examination of the chess relations, it appears that subjects were not attending to the attack relation as much as the defense relation. But a casual look at M's protocols indicated that some attacking pieces were clustered in his protocols.

Therefore, we analyzed the 14 middle-game and puzzle positions to find the strongest attacks; 18 such attacks were found, consisting mostly of pieces attacking the opponent's king position. Of these 18 attacks, 11 were chunked in M's protocols, in the sense that at least two of the attacking pieces appeared within the same chunk; rarely did the attacked pieces also appear in the same chunk with the attackers.

Thus, it appears that two kinds of attacks get chunked. The first kind is a fortuitous attack characterized by an attack relation between two adjacent pieces. The second is more abstract and involves combinations of pieces of the same color converging, usually, on the opponent's king position. These attack chunks would also be stereotyped, often involving classic maneuvers against a stereotyped defensive position. M would be able to recognize all these chunks provided that he has stored in long-term memory a modest vocabulary of variant patterns for each of a half dozen types of configurations.

Thus, we can account for M's performance in recalling positions he has seen for 5-seconds if we postulate that he has a short-term memory of average capacity, but a long-term memory capable of recognizing: a variety of chunks consisting of pawns; a variety of chunks consisting of common first-rank configurations; a variety of chunks consisting of common pawn chain, rook pair, and rook and queen configurations; a variety of common configurations of attacking pieces, especially along a file, diagonal, or around an opponent's castled-king position.

By confronting chess players of varying strength with a perception task and a memory task, we have shown that the amount of information extracted from a briefly exposed position varies with playing strength.

By measuring the time intervals between placements of successive pieces when the subjects attempted to reconstruct the positions, we were able to identify the boundaries of perceptual chunks. The data suggest that the superior performance of stronger players derives from the ability of those players to encode the position into larger perceptual chunks, each consisting of a familiar subconfiguration of pieces. Pieces within a single chunk are bound by relations of mutual defense, proximity, attack over small distances, and common color and type.

Condensed with permission from *Cognitive Psychology*, 4(1):55-81, 1973. Reprints of original article from William G. Chase, Ph.D., Department of Psychology, Carnegie-Mellon University, Pittsburgh, Pa. 15213