

Perception in Chess¹

WILLIAM G. CHASE AND HERBERT A. SIMON

Carnegie-Mellon University

This paper develops a technique for isolating and studying the perceptual structures that chess players perceive. Three chess players of varying strength — from master to novice — were confronted with two tasks: (1) A perception task, where the player reproduces a chess position in plain view, and (2) de Groot's (1965) short-term recall task, where the player reproduces a chess position after viewing it for 5 sec. The successive glances at the position in the perceptual task and long pauses in the memory task were used to segment the structures in the reconstruction protocol. The size and nature of these structures were then analyzed as a function of chess skill.

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 (Simon & Barenfeld, 1969). But we know from other considerations that he is seeing much more. Our work is concerned with just what the expert chess player perceives.

The most extensive work to date on perception in chess is that done by de Groot and his colleagues (de Groot, 1965, 1966; Jongman, 1968). In his search for differences between masters and weaker players, de Groot was unable to find any gross differences in the statistics of their thought processes: the number of moves considered, search heuristics, depth of search, and so on. Masters search through about the same number of possibilities as weaker players—perhaps even fewer, almost certainly not more—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.

De Groot did, however, find an intriguing difference between masters and weaker players in his short-term memory experiments. Masters showed a remarkable ability to reconstruct a chess position almost perfectly after viewing it for only 5 sec. There was a sharp dropoff in

¹ Correspondence should be addressed to: Dr. William G. Chase, Department of Psychology, Carnegie-Mellon University, Pittsburgh, PA 15213. This work was supported by Public Health Service Research Grant MH-07722 from the National Institute of Mental Health. We are indebted to Hans Berliner for his masterful performance as a subject.

this ability for players below the master level. This result could not be attributed to the masters' generally superior memory ability, for when chess positions were constructed by placing the same numbers of pieces randomly on the board, the masters could then do no better in reconstructing them than weaker players. Hence, the masters appear to be constrained by the same severe short-term memory limits as everyone else (Miller, 1956), and 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.

One key to understanding chess mastery, then, seems to lie in the immediate perceptual processing, for it is here that the game is structured, and it is here in the static analysis that the good moves are generated for subsequent processing. Behind this perceptual analysis, as with all skills (cf., Fitts & Posner, 1967), lies an extensive cognitive apparatus amassed through years of constant practice. What was once accomplished by slow, conscious deductive reasoning is now arrived at by fast, unconscious perceptual processing. It is no mistake of language for the chess master to say that he "sees" the right move; and it is for good reason that students of complex problem solving are interested in perceptual processes (cf., Newell & Simon, 1972). Our main concern here is to discover and characterize the structures, or chunks, that are seen on the board and stored in short-term memory.

SCOPE OF THE STUDY

The previous studies of chess perception make highly plausible the hypothesis that the chess master encodes information about a position in chunks, but provides no direct methods for delimiting the chunk boundaries or detecting the 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 player's perceptual processing of the board is so rapid (and probably unavailable to conscious introspection) that it is impossible to obtain an accurate verbal description of the process from him. Although eye movements give us a record of how the board is scanned (de Groot, 1966; Simon & Barenfeld, 1966; Tichomirov & Poznyanskaya, 1966; Winikoff, 1967), they don't tell us precisely which pieces are observed (especially in peripheral vision) and in what order; they only

tell us the general area being aimed at by the fovea. And, of course, data on eye movements can't tell us what information is being abstracted from the display.

There are, however, other techniques, which have been used with verbal materials, that would appear promising for the problem at hand. Tulving (1962) has looked at clusters in free recall protocols; and Bower and Springston (1970) have looked at the timing relations and pauses in the output. McLean and Gregg (1967) have used pauses to define chunks in rote learning. Ein-Dor (1971) has studied chunking of visual stimuli in the form of Chinese ideograms, using a method essentially identical with our perception experiment.

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. In the *perception task*, we ask chess players to reconstruct a chess position while it remains in plain view, and we use subjects' successive glances at the board as an index of chunking. The basic assumption is that, under the conditions of the experiment, the subject will encode only one chunk per glance while reconstructing the position.

In the *memory task*, which is very similar to de Groot's task, we ask chess players to reconstruct a position from memory after brief exposure to it, and we use the timing or clustering in recall to segment the output into chunks.

The memory task permits us to replicate the basic findings of de Groot and Jongman. These results are so important that it is essential to have an independent replication; moreover, the empirical results for the case of the random boards have never been reported in detail in the literature.

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.

In the following sections of this paper, we will report and analyze the main body of data obtained by presenting the two tasks to a chess master and to weaker players. Then we will investigate in somewhat greater detail the data for the chess master in middle game positions. In a final section, we will summarize our findings and our interpretation of them.

METHOD

Three chess players, a master (M), a Class A player (A), and a beginner (B), were used as subjects. Twenty games were selected from chess books and magazines to generate the stimuli. These were games

between advanced players (masters, experts, and perhaps a few Class A players). 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 the 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 in.: 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 procedure in the memory task was similar to that used by de Groot (1965), except that the subject was given multiple trials in each position. 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 sec, 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 min). 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 sec 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 the perception task, each subject processed five middle-game positions, five end-game positions, two randomized middle-game positions, and two randomized end-game positions. He also processed the same number of each kind of position in the memory task.

RESULTS

The videotape records for both tasks were analyzed by recording each piece as it was placed on the board, and by recording the time, within $\frac{1}{10}$ sec, 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: (1) *attack*: either one of the two pieces attacks the other; (2) *defense*: either one of the two pieces defends the other; (3) *proximity*: each piece stands on one of the eight squares adjacent to the other; (4) *common color*: both pieces are of the same color; and (5) *common type*: both pieces are of the same type (e.g., both are pawns, rooks, etc.).

Accuracy of Reconstruction

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

Figure 1 shows the results for the middle-game positions, actual and random. Figure 2 shows the results for the end-game positions, actual and random. The figures show the average number of pieces placed correctly by each subject on successive trials for all positions of the type in question. The standard errors, based on five scores, are shown for the first trial of the middle- and end-game positions.

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 was able to reproduce 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 (as many as 14 in one case). 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. Of course, A and B, because

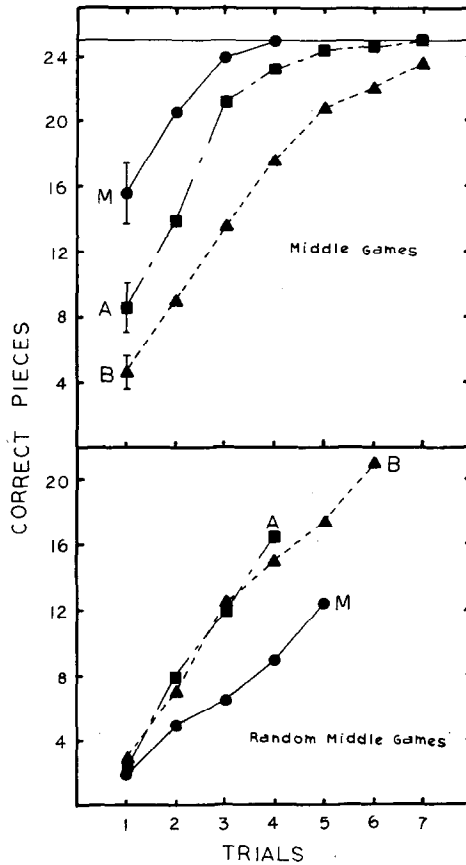


FIG. 1. Learning curves of the master (M), class A player (A), and beginner (B) for the middle-game and random middle-game positions. The brackets are standard errors on five positions.

of their poorer first-trial performance, had much more room for improvement than did M; this difference disappears 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-sec 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

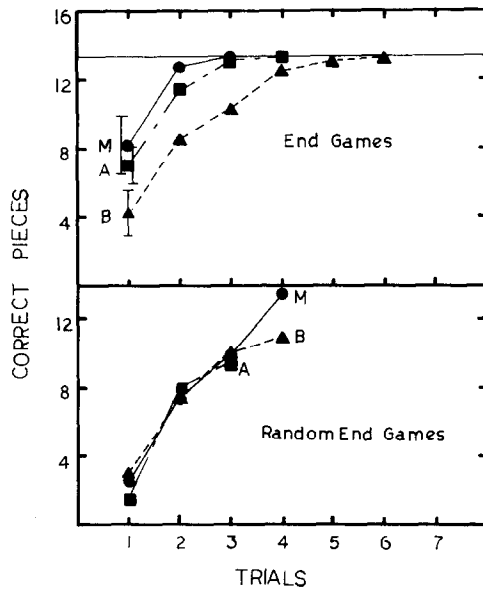


FIG. 2. Learning curves for the end-game and random end-game positions. The brackets are standard errors based on five positions.

first-trial performances of all three subjects on the random positions was even *poorer* than B's performance on the actual game positions.

There is some quantitative difference between M's performance on the actual middle-game positions and the performance reported by de Groot for grandmasters and masters in middle-game positions. Typically, de Groot's grandmaster and master subjects were able to replace about 23 or 24 pieces out of 25 correctly after 5 sec (or less!) view of the board. M, as we have seen, averaged only 16 pieces. The most plausible explanation for the difference lies in the nature of the positions used in the tests. De Groot used positions from relatively recent grandmaster games (not known to the subjects), and excluded positions that were not "quiet," i.e., positions where exchanges of pieces were in midstream. Some of the positions we used were from games between players of less than master caliber, and in several of them exchanges were under way.

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 (Reinfeld, 1945). 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-sec trial was conducted. For these new positions, B, A, and M averaged 33, 49, and 81% correct, respectively, as compared to 18, 34, and 62%, respectively, on the first trial of the previous positions. These figures are in very close agreement with those published by de Groot (1966), and, taking the differences in stimuli into account,² our data unequivocally replicate de Groot's important results.

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%, 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: (1) *within-glance intervals*, intervals between pieces placed without looking back at the original position; and (2) *between-glance intervals*, intervals between two pieces separated by a glance back at the original position. These frequency distributions are shown in Fig. 3 for each subject.

The results are straightforward and roughly the same, with one exception, for all three subjects. Within-glance intervals seldom exceeded 2 sec, and the modal intervals were $\frac{1}{2}$ sec or less. For the between-glance intervals, there was a tendency for the better players to take less time: the mean latencies were 2.8, 3.2, and 3.5 sec for M, A, and B, respectively. The differences between these means are statically significant ($p < .05$) when tested against a pooled error term.

In the memory task, of course, 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

²There were other differences between de Groot's procedure and ours. For example, de Groot always informed his subjects about who was on move (white or black), and the subject always viewed the board from that perspective, whereas our subjects didn't know who was on move and they always viewed the board from the perspective of the white player. These differences would seem to be minor, however, compared to the differences in quiet positions.

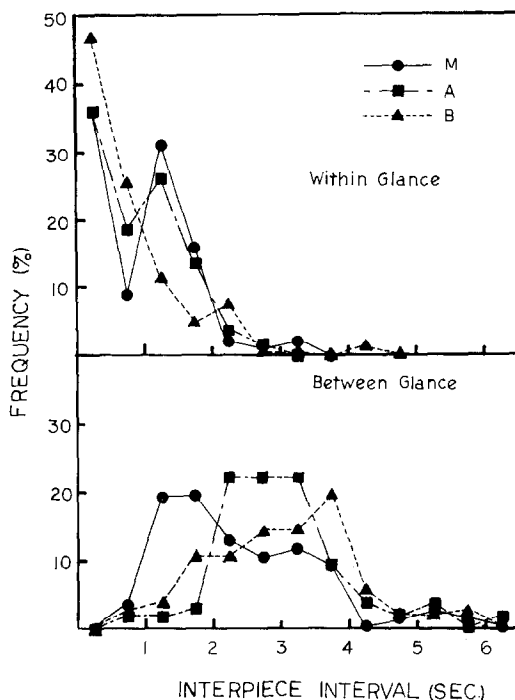


FIG. 3. The frequency distributions, for each subject, of the interpiece intervals for the within glance and between glance times of the perception task.

combined frequency distribution for the perception task. Figure 4 compares the combined distributions for each subject in the perception task with the trial 1 distributions of the memory 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 sec but a "tail" of longer intervals. In the memory task, there are numerous intervals up to about 2½ sec, and again a tail of longer intervals. The very short intervals in the distributions, ½ sec 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 sec to retrieve a piece from the side of the board.³

³ A possible artifact in this experiment is the time required actually to pick up the pieces. We have replicated the experiment as an oral task (pointing to squares and naming the pieces on them) and a paper-and-pencil task (writing the names of the pieces on the proper squares of a diagram on the board) without altering the results. We will report on these later experiments in a subsequent paper.

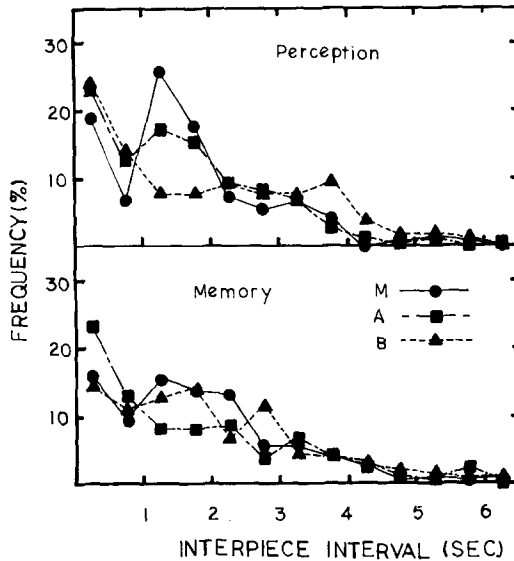


FIG. 4. A comparison of the frequency distributions of the interpiece intervals for the perception and memory experiments.

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

1. The pieces placed on the board by the subject in the perception task after a single glance correspond to a single chunk. About 2 sec 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.

2. A sequence of pieces placed on the board by the subject in the memory task with intervals of less than 2 sec 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.

The hypothesis gains some plausibility from measurements in previous experiments of the times required to transfer information into short-term memory. In particular, Dansereau (1969), studying times of performance of mental arithmetic and related tasks, estimated that about

2 sec was needed to begin processing a chunk whose label was held in short-term memory, and only about 300 msec to transfer to short-term memory each successive element of the chunk. Intervals even shorter than 300 msec intervals are familiar from other experiments on the speed with which subjects can count down familiar lists (Landauer, 1962; Pierce & Karlin, 1957).

If our hypothesis is correct (that time intervals of 2 sec 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 (none of our results would be essentially changed if we had adopted a 2½ instead of a 2-sec boundary). We turn next to these tests of the hypothesis.

Chess Relations: Perception Task

Table 1 shows, for each subject, the within-glance probabilities and mean interpiece latencies for each of the 16 possible combinations of attack (A), defense (D), same color (C), same piece (S), and proximity (P) relations. For example, the first row shows latencies and probabilities for successive pieces which have no relation; they are of opposite color, are not proximate to each other, and are not of the same type. The second row is for pieces that have only an attack relation. The last row (DPCS) is for pieces that have a defense relation, are within one square of each other, are of the same color, and are of the same type. This last row is comprised almost totally of pawn chains. Notice also that the color relation is carried redundantly with the defense relation (pieces defending one another are of the same color). Table 2 shows the corresponding data for between-glance probabilities and latencies.

The first thing to notice is that these data are quite similar for all subjects. The latencies show the same systematic trends, and, for the probabilities, the product moment correlation 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, shown in Table 2, 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.

The marginal row statistics, shown in the last six columns of Tables 1

TABLE 1
Chess Relations: Within-Glance Latencies (RT) and Probabilities (p) for Each Subject (M,A,B), and Total Frequencies (N), Average Latencies (RT), Standard Error of Average Latencies (SE_{RT}), Observed Probabilities (P_o), a Priori Probabilities (P_e), and Deviation Scores (Z) for Combinations of the Five Chess Relations: Attack (A), Defense (D), Spatial Proximity (P), Same Color (C), and Same Piece (S)

Relations	M			A			B			All players				
	RT	p		RT	p		RT	p		$\overline{RT}$	$SE_{\overline{RT}}$	P_o	P_e	Z
—	4.63	.044		1.48	.035		—	—	0	3.05	1.321	.028	.320	-30.1
A	1.60	.011		1.90	.009		1.40	.013	.013	1.63	.119	.010	.0201	-1.6
P	1.30	.011		1.80	.009		—	0	0	1.55	.177	.007	.0057	.3
C	1.53	.099		1.40	.060		1.75	.125	.125	1.58	.173	.091	.255	-9.7
S	.80	.022		1.35	.017		.80	.013	.013	1.02	.163	.017	.148	-16.9
AP	1.23	.044		1.85	.052		1.78	.063	.063	1.65	.206	.052	.0077	3.4
AS	—	0		—	0		—	0	0	—	—	0	.0025	—
DC	1.53	.099		1.44	.103		1.38	.100	.100	1.45	.104	.101	.0423	3.3
PC	1.17	.033		1.51	.060		.72	.063	.063	1.18	.126	.052	.0159	2.8
PS	2.10	.022		.77	.026		.30	.013	.013	1.13	.322	.021	.0075	1.6
CS	1.44	.077		1.08	.043		.58	.050	.050	1.11	.190	.056	.0939	-2.8
APS	—	0		—	0		—	0	0	—	—	0	.0022	—
DPC	1.30	.132		1.19	.190		1.04	.113	.113	1.19	.081	.150	.0469	4.9
DCS	1.50	.044		.50	.017		—	0	0	1.17	.385	.021	.0057	1.8
PCS	.41	.154		.46	.155		.41	.188	.188	.43	.043	.164	.0105	7.0
DPCS	.53	.209		.41	.353		.48	.263	.263	.47	.033	.230	.0162	8.6
	1.22			.99			.89			1.04				

TABLE 2
 Chess Relations: Between-Glance Latencies (RT) and Probabilities (p) for Each Subject (M,A,B), and Total Frequencies (N), Average Latencies ($\overline{RT}$), Standard Error of Average Latencies ($SE_{\overline{RT}}$), Observed Probabilities (P_o), *a Priori* Probabilities (P_e), and Deviation Scores (Z) for Combinations of the Five Chess Relations: Attack (A), Defense (D), Spatial Proximity (P), Same Color (C), and Same Piece (S)

Relations	M			A			B			All players					
	RT	p		RT	p		RT	p		$\overline{RT}$	$SE_{\overline{RT}}$	P_o	P_e	Z	
—	3.76	.213		3.25	.255		5.20	.156		4.00	.434	.203	.320	-4.2	
A	2.50	.013		3.80	.055		3.63	.052		3.55	.926	.038	.0201	1.4	
P	2.35	.025		4.60	.036		3.52	.065		3.50	.527	.042	.0057	2.7	
C	2.94	.338		3.15	.236		3.38	.260		3.14	.182	.283	.255	.9	
S	2.63	.050		2.75	.036		3.43	.052		2.97	.311	.047	.148	-6.9	
AP	1.25	.025		3.05	.036		3.70	.052		2.92	.638	.038	.0077	2.3	
AS	—	0		—	0		—	0		—	0	0	.0025	—	
DC	1.63	.050		2.91	.127		3.33	.156		2.91	.277	.108	.0423	3.1	
PC	1.27	.038		1.20	.018		3.00	.026		1.83	.346	.028	.0159	1.1	
PS	4.63	.038		3.10	.018		—	0		4.25	1.121	.019	.0075	1.2	
CS	4.06	.063		8.10	.055		2.75	.026		5.01	1.435	.047	.0939	3.2	
APS	—	0		—	0		—	0		—	0	0	.0022	—	
DPC	1.55	.138		2.58	.073		2.51	.117		2.08	.157	.113	.0469	3.0	
DCS	—	0		.90	.036		—	0		.90	0	.009	.0057	.6	
PCS	3.50	.013		—	0		5.20	.026		4.63	.484	.014	.0105	.5	
DPCS	—	0		1.90	.018		.70	.013		1.30	.424	.009	.0162	-1.0	
	2.86			3.30			3.58			3.24					

and 2 are, therefore, representative of all subjects. The first summary column shows the total frequency for each type of event. The second and third columns show the mean and standard error of the interpiece latencies. The fourth column shows the probabilities, based on the frequencies of the first column. The fifth column shows the *a priori* probabilities, which would prevail if successive pieces were chosen at random.⁴ The last column shows a deviation score (the observed probability minus the *a priori* probability, divided by the standard error) assuming the normal approximation to the binomial. The *a priori* values can be considered exact since they are based on about 12,000 observations.

A comparison of Tables 1 and 2 reveals quite different patterns for the within-glance and between-glance probabilities. An examination of the *z* scores shows that 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. As mentioned before, 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. Note also that these two relations 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, of course, also implies the same-color relation. The low frequencies for the A relation suggests that attacks are noticed only if the pieces are in close spatial proximity (but see later the additional comments on this point in the discussion of the Perceptual Chunks of a Chess Master). 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

⁴ The *a priori* probabilities were calculated by first recording, for each position, all relations that exist between every possible pair of pieces; the *a priori* probability for a relation, then, is simply the total number of occurrences of a relation divided by the total number of possible pairs. The *a priori* probabilities were based on 30 positions, and the random *a priori* probabilities were based on eight random positions.

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. For example, the DPC, CS, and DC relations are slightly higher and the S and null relations are down somewhat from the chance level. 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.

Table 3 shows the summary data for the between- and within-glance data for the randomized positions. Although there weren't many observations on individual subjects, the pattern of probabilities was still the same across subjects. The interesting thing about these data is that they look very similar to the data from real positions. Notice that 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 procedure of the perception experiment offers no absolute guarantee that the subject did not pick up more than one chunk at a glance. However, subjects reported that it was most comfortable to glance frequently at the board and not to retain much information in short-term memory. Moreover, especially with M, there was no evidence of perseveration in glances. The duration of most of his glances, including time for the head movement and time to place the next piece, was close to the 2-sec boundary, and almost none was more than 4 sec long. But the main test of the one-glance-one-chunk hypothesis lies in comparison of the data between perception and memory experiments.

Chess Relations: Memory Task

Table 4 shows the memory data for individual subjects.⁵ Again the patterns of latencies and probabilities look the same 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? The evidence seems fairly good on this point. It can be seen

⁵ Only the chess relations from actual game positions were analyzed for the memory task. Trial 1 recall of the random positions was so poor that there simply weren't enough data to make any comparisons.

TABLE 4
 Chess Relations for the Memory Experiment: Coverage Latencies ($\overline{RT}$), Observed Probabilities (P_o), Frequency (N), Standard Error of Average Latencies ($SE_{\overline{RT}}$), *a Priori* Probabilities (P_a), and Deviation Scores (Z) for Combinations of the Five Chess Relations: Attack (A), Defense (D), Spatial Proximity (P), Same Color (C), and Same Piece (S)

Relations	M			A			B			All players				
	$\overline{RT}$	P_o	$\overline{RT}$	P_o	$\overline{RT}$	P_o	$\overline{RT}$	P_o	N	$\overline{RT}$	$SE_{\overline{RT}}$	P_o	P_e	Z
—	6.67	.147	5.18	.117	4.53	.124	114	5.81	.552	.132	.320	—16.4		
A	2.86	.012	4.94	.019	3.60	.010	12	3.85	1.026	.014	.0201	—1.6		
P	2.53	.010	3.20	.004	8.00	.005	6	3.55	.886	.007	.0057	.4		
C	3.57	.150	6.28	.140	3.04	.163	129	4.19	.426	.149	.255	—8.7		
S	3.24	.065	7.35	.039	1.22	.025	41	4.00	.726	.047	.148	—13.9		
AP	2.80	.007	2.80	.012	5.00	.025	16	3.18	.552	.018	.0077	2.4		
AS	2.40	.010	1.60	.012	1.80	.005	8	2.02	.289	.009	.0025	2.1		
DC	3.99	.070	2.96	.035	3.18	.064	50	3.60	.627	.038	.0423	2.0		
PC	3.37	.027	6.10	.019	2.90	.045	25	3.75	.840	.029	.0159	2.3		
PS	.40	.002	.80	.008	4.60	.010	5	2.24	1.135	.006	.0075	—7		
CS	2.86	.055	4.83	.027	1.90	.054	40	2.94	.585	.046	.0939	—6.7		
APS	.50	.005	—	0	—	0	2	.50	0	.002	.0022	.1		
DPC	2.22	.175	2.69	.214	3.01	.213	168	2.58	.207	.194	.0469	11.0		
DCS	1.19	.035	.65	.078	2.28	.050	44	1.19	.128	.031	.0057	6.0		
PCS	.63	.107	.92	.167	.67	.114	109	.75	.116	.126	.0105	10.2		
DPCS	.97	.122	1.08	.109	.73	.094	96	.95	.111	.111	.0167	8.9		
	2.98		3.19		2.68			2.97						

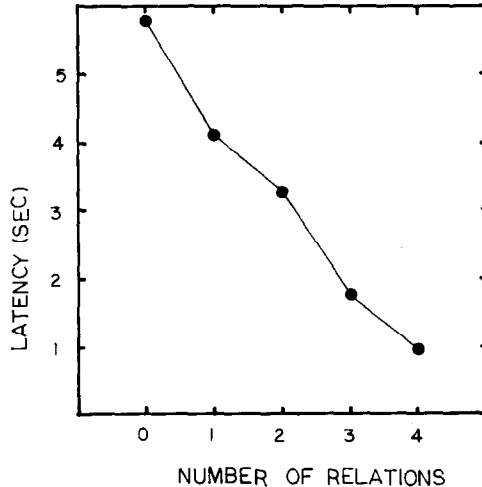


FIG. 5. The relation between interpiece latencies and the number of relations between pairs of successively placed pieces in the memory task.

in Table 4 that longer latencies are associated with fewer interpiece chess relations, and Fig. 5 illustrates the relation between average interpiece latencies and the number of chess relations between the pieces. Another indication of this relationship is that latencies are correlated $-.73$ with the z scores of $(P_o - P_e)/SE_{P_o}$.

We should also note in passing that errors usually occur toward the end of the protocol; subjects usually report first what they know, and fast, and these errors generally have long latencies and few relations. Also, the results remain the same if we score only correctly placed pieces.

A closer look at Table 4 reveals that the lowest latencies (except for APS, but we will not consider it because it occurred only twice, and both of those for M) 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 were 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.

We next turn to the hypothesis that time intervals of roughly 2 sec or more correspond to boundaries between chunks. If this hypothesis is correct, the chess relations with latencies greater than 2 sec ought to look like chance occurrences, whereas the relations occurring within 2 sec ought to show even more structure. Table 5 shows the memory data of Table 4 partitioned into relations for latencies less than (or

equal to) 2 sec, and chess relations for latencies greater than 2 sec. It is clear from Table 5 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 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 we wish to consider 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, as in Table 6, and contrast them with the *a priori*

TABLE 6

A Comparison of the Perceptual, Memory, and *a Priori* Chess Relation Probabilities for Combinations of the Five Chess Relations: Attack (A), Defense (D), Spatial Proximity (P), Same Color (C), and Same Piece (S)

Chess relations	Perception		Memory	Perception		Memory		
	Within-glance (random)	Within-glance (games)	Less than 2 sec	Between-glance (random)	Between-glance (games)	Greater than 2 sec	<i>A priori</i> (games)	<i>A priori</i> (random)
—	.018	.028	.031	.214	.203	.258	.320	.298
A	.009	.010	.010	.009	.038	.018	.0201	.0265
P	.018	.007	.004	.027	.042	.010	.0057	.0046
C	.063	.091	.089	.232	.283	.224	.255	.287
S	.018	.017	.029	.098	.047	.070	.148	.150
AP	.143	.052	.015	.080	.038	.023	.0077	.0351
AS	.018	0	.010	.018	0	.008	.0025	.0013
DC	.054	.101	.054	.054	.108	.063	.0423	.0238
PC	.089	.052	.027	.063	.028	.031	.0159	.0126
PS	.089	.021	.008	.027	.019	.003	.0075	.0139
CS	.054	.056	.046	.054	.047	.047	.0939	.101
APS	0	0	.004	0	0	0	.0022	.0053
DPC	.170	.150	.198	.107	.113	.190	.0469	.0265
DCS	.054	.021	.079	.054	.009	.016	.0057	.0013
PCS	.179	.164	.216	.018	.014	.013	.0105	.0060
DPCS	.027	.230	.179	.027	.009	.026	.0162	.0066

TABLE 7
Intercorrelation Matrix for the Perceptual, Memory, and *a Priori* Chess
Relation Probabilities

		1	2	3	4	5	6	7
1 Within-glance (random)	} Perception							
2 Within-glance (games)		.49	.59	.06	.02	.09	—	.19
3 Less than 2 sec	} Memory		.89	.06	.12	.18	—	.04
4 Between-glance (random)				.08	.10	.23	—	.03
5 Between-glance (games)	} Perception				.92	.93		.91
6 Greater than 2 sec						.91		.81
7 <i>A priori</i>								.87

probabilities. There are some slight differences between the perceptual and memory probabilities, but these differences are everywhere small compared to their differences with the *a priori* probabilities. Table 7 illustrates this similarity more sharply by showing the matrix of correlations derived from Table 6. There are two clusters of correlations in this table. 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 *a priori* 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 sec between the placing of two pieces appear to be significantly different from the processes that occur during an interval of less than 2 sec. 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. Recall that in the perception task there was a systematic difference in the duration of the glances as a function of chess skill, with less time being taken by the more skilled players. But the average number of pieces per glance did not vary systematically as a function of chess skill. For the middle-game positions, the average number of pieces per glance was 2.0, 2.8, and 2.0, respectively, for M, A, and B. For the memory experiment, however, the corresponding number

of pieces per chunk was 2.5, 2.1, and 1.9, respectively. Thus, it appears that the chunks are about the same size in both 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.

We undertake next to examine further evidence that will help us decide whether the chunks defined by long pauses have the properties we would expect from our previous experimental knowledge of perceptual chunking.

CHUNK SIZE AND MEMORY SPAN

Having segmented the recall protocol into chunks, we are now in a position to 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 general practice of the subjects was to place first those groups of pieces they thought they remembered well, then to search memory for additional pieces. During the first phase, placing pieces in recall without "problem solving," chunks tended to be relatively large and errors relatively few. During the second phase, pieces tended to be placed one by one (pawns sometimes by pairs or triads), 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 (i.e., where they would function well, or where they are often posted in actual games), rather than recalling where he had actually seen them. This behavior was more true of M than the other subjects. De Groot (1966) points out, in fact, that subjects can average better than 44% simply by putting down the "average" or prototype position derived from master games.

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 sec or more) followed by mostly errors, or a series of long pauses (5 sec or more) with errors. Based on this criterion, Table 8 shows, for each of the subjects in the memory experiment, the average sizes of eight successive chunks on the first trial for the actual middle-game and end-game positions. The last column of the table shows the average number of chunks recalled for the first trial in each of these positions.

TABLE 8
Average Sizes of Successive Chunks for Each Player, Middle-Game and End-Game Positions. Memory Experiments, First Trial

		Successive chunks								Average chunks/ trial
		1	2	3	4	5	6	7	8	
Middle games	M	3.8	3.0	2.5	2.3	1.9	1.5	2.2	2.0	7.7
	A	2.6	2.5	1.8	1.6	1.7	1.7	2.1	2.5	5.7
	B	2.4	2.1	2.0	1.6	1.4	1.5	1.0	2.0	5.3
End games	M	2.6	1.6	1.4	1.8	1.8	1.2	2.3	1.0	7.6
	A	2.4	1.4	2.0	2.0	1.0	1.0	1.0	1.0	6.4
	B	2.2	2.4	2.2	1.0	1.0	1.0	1.0	0	4.2

We observe, first, that 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. First, it may be that subjects simply recall their larger chunks first. Second, it is well known that recall has an interfering effect on short-term memory, and it may be that interference causes large chunks to break up into smaller chunks as some of the relations are forgotten. Third, the 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.

We observe, second, that 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. This seems a rather surprising result; we know 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. First, it may be that M does not store a small number of unrelated chunks in short-term memory. Rather, he may be able to organize the chunks on the board in some as yet undertermined way so that more chunks can be stored. In this way, M will get more information from the partially reconstructed board than weaker players about what the rest of the position should be. In other words, the data should make us skeptical of an overly simple theoretical position that postulates that

short-term memory consists of a linear list of seven or so unrelated chunk slots.

A second possibility, discussed earlier, 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

De Groot and Jongman have made some observations on the nature of the perceptual chunks into which grandmasters and masters encode information. In their experiments, however, these authors had no objective means for detecting chunk boundaries. Our data give us an operational method of characterizing chunks, which we will apply to the middle-game memory experiments of subject M.

Table 9 shows for M the sizes of successive chunks for the five middle-

TABLE 9
Size of the Master's Successive Chunks for the Five Middle-Game and 9 Puzzle Positions. Memory Experiment, First Trial

	Successive chunks												
	1	2	3	4	5	6	7	8	9	10	11	12	13
M1	6	7	1	2	2	2	1	3	1	1			
M2	3	2	1	2	2	1	3	1	1	2	2	1	2
M3	4	2	5	2	4	1	1	1	2				
M4	4	2	1	1	1	1	5	1	2				
M5	2	2	2	4									
P1	3	7	4	2	1	2	1	2	2				
P2	4	5	1	2	1	3							
P3	2	2	2	4	3	3	2						
P4	6	3	5	2	1								
P5	5	2	3	3	2	1	1	1	1	1	1		
P6	3	4	4	1	1	1	4	4					
P7	8	4	5	3	3								
P8	2	1	2	4	4	1	2	2	2				
P9	3	2	1	2	2	2	2	3	2	3	3	1	
	3.9	3.2	2.6	2.6	2.1	1.6	2.2	2.0	1.6	1.8	2.0	1.0	2.0

SE = .46.

game and nine puzzle positions for trial 1 of the memory experiments. The great bulk of the 77 chunks (two or more pieces within 2 sec) 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. Over half the chunks (47) contained a Pawn chain, sometimes with nearby supporting pieces and sometimes with blockading pieces or contain more than one of these categories. For example, a castled-King position (a strong and often-used defensive structure) sometimes with nearby pieces. Twenty-seven chunks consisted of clusters of pieces of the same color (exclusive of castled-King positions), and 18 of these were of very familiar types: nine chunks consisted of pieces on the back rank (rank 1 or 8), often in their original undeveloped position; and nine chunks consisted of connected Rooks (mutually supporting), or the Queen connected with one or two Rooks — a very powerful attacking structure. These categories are not mutually exclusive since some chunks contain more than one of these categories. For example, a castled-King position also contains a Pawn chain, and sometimes Pawn chains and clusters of pieces occur within the same chunk. The point is, however, that over 75% 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. Recall that the attack relation appeared more often than chance only if the attacking piece was also on an adjacent square. But a casual look at M's protocols indicated that some attacking pieces were clustered in his protocols.

Therefore, to test this hypothesis more objectively, the 14 middle-game and puzzle positions were analyzed by the authors 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. Of the 11 attacks, six consisted of Rook and Queen-Rook combinations—one chunk also contained a Pawn in combination with the Queen and Rook, and the other five chunks consisted of a Knight in combination with a Queen or Rook.

Thus, it appears that there are two kinds of attacks that get chunked. The first kind is a fortuitous attack characterized by an attack relation between two adjacent pieces. The second kind of attack is more abstract and involves combinations of pieces of the same color converging, usually, on the opponent's King position. The relation between the at-

tacking pieces wouldn't appear as an attack relation; these pieces would either have no relation or a defense relation. 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. The estimates given in Simon and Barenfeld (1969) as to the size of vocabulary required appear now to be, if anything, somewhat too large.

Thus, we can account for M's performance in recalling positions he has seen for 5 sec if we postulate that he has a short-term memory of average capacity, but a long-term memory capable of recognizing:

1. A variety of chunks consisting of Pawns (and possibly Rook and minor pieces) in common castled-King configurations;
2. A variety of chunks consisting of common first-rank configurations;
3. A variety of chunks consisting of common Pawn chain, Rook pair, and Rook and Queen configurations;
4. A variety of common configurations of attacking pieces, especially along a file, diagonal, or around an opponent's castled-King position.⁶

CONCLUSION

By confronting chess players of varying strength, from master to novice, 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, thus confirming earlier experiments of de Groot, Jongman, and others.

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 (which does not appear in random positions) 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.

There is also some evidence that chunks may be held together by more abstract relations. There are more chunks in recall for the stronger players, yet the frequencies of between-chunk relations (of the kinds we recorded) are all close to chance. This may derive from a hierarchical

⁶The master's vocabulary of recognizable configurations inferred by Jongman (1968) is very similar to the list above.

organization of the chunks, related to chess skill, that is more abstract than the simple chess relations we have measured. Further, in M's protocol there is good evidence that pieces converging on the opponent's King position (or sometimes on other vulnerable positions) are chunked—a more abstract but fairly well-defined attack relation.

Finally, the number of chunks retained in short-term memory after brief exposure to chess positions is about of the magnitude we would predict from immediate recall of common words (Miller, 1956) and copying of visual patterns (Ein-Dor, 1971).

REFERENCES

- BOWER, G. H., & SPRINGSTON, F. Pauses as recoding points in letter series. *Journal of Experimental Psychology*, 1970, **83**, 421-430.
- DANSEREAU, D. An information processing model of mental multiplication. Unpublished doctoral dissertation. Carnegie-Mellon University, 1969.
- DE GROOT, A. D. *Thought and choice in chess*. The Hague: Mouton, 1965.
- DE GROOT, A. D. Perception and memory versus thought: Some old ideas and recent findings. In B. Kleinmuntz (Ed.) *Problem solving*. New York: Wiley, 1966.
- EIN-DOR, P. Elements of a theory of visual information processing. Unpublished doctoral dissertation. Carnegie-Mellon University, 1971.
- FITTS, P. M., & POSNER, M. I. *Human Performance*. Belmont, CA: Brooks/Cole, 1967.
- JONGMAN, R. W. *Het oog van de Meester*. Amsterdam: van Gorcum, 1968.
- LANDAUER, T. K. Rate of implicit speech. *Perceptual and Motor Skills*, 1962, **15**, 646.
- MCLEAN, R. S., & GREGG, L. W. Effects of induced chunking on temporal aspects of serial recitation. *Journal of Experimental Psychology*, 1967, **74**, 455-459.
- MILLER, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 1956, **63**, 81-97.
- NEWELL, A., & SIMON, H. A. *Human problem solving*. New York: Prentice-Hall, 1972.
- PIERCE, J. R., & KARLIN, J. E. Reading rates and the information rate of a human channel. *Bell Telephone System Technical Publications*, 1957, **36**, 497-516.
- REINFELD, F. *Win at chess*. New York: Dover, 1945.
- SIMON, H. A., & BARENFELD, M. Information processing analysis of perceptual processes in problem solving. *Psychological Review*, 1969, **76**, 473-483.
- TICHOMIROV, O. K., & POZNYANSKAYA, E. D. An investigation of visual search as a means of analyzing heuristics. *Soviet Psychology*, Winter 1966-67, **5**, 2-15.
- TULVING, E. Subjective organization in free recall of "unrelated" words. *Psychological Review*, 1962, **69**, 344-354.
- WINIKOFF, A. W. Eye movements as an aid to protocol analysis of problem solving behavior. Unpublished doctoral dissertation. Carnegie-Mellon University, 1967.

(Accepted April 7, 1972)