

Gobet, F. & Simon, H. A. (1996). Recall of random and distorted positions: Implications for the theory of expertise. *Memory & Cognition*, 24, 493-503.

Recall of rapidly presented random chess positions is a function of skill

Fernand Gobet and Herbert A. Simon

Department of Psychology
Carnegie Mellon University

Correspondence to:

Prof. Herbert A. Simon
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213

has@cs.cmu.edu
412-268-2787

Abstract

A widely cited result asserts that experts' superiority over novices in recalling meaningful material from their domain of expertise vanishes when random material is used. A review of recent chess experiments where random positions served as control material (presentation time between 3 and 10 seconds) shows, however, that strong players generally maintain some superiority over weak players even with random positions, although the relative difference between skill levels is much smaller than with game positions. The implications of this finding for expertise in chess are discussed and the question of the recall of random material in other domains is raised.

Recall of rapidly presented random chess positions is a function of skill

A classical result in the study of expertise is that experts are better than non-experts at memorizing meaningful material from their domain of expertise, but lose their superiority when the material is randomized. This result was first obtained in the study of chess memory (Chase and Simon, 1973a; Jongman and Lemmens, cited in Vicente & De Groot, 1990), and has since been widely cited in cognitive psychology textbooks (e.g. Anderson, 1990; Lesgold, 1988), and hailed as one of the cornerstones of the study of expertise (Saariluoma, 1989). The basic relation between skill and meaningfulness has been replicated in various domains, although wide variations in the presentation time of the stimuli make quantitative comparisons difficult: Go (Reitman, 1976); bridge (Engle & Bukstel, 1978; Charness, 1979); Othello (Wolff, Mitchell & Frey, 1984); electronics (Egan & Schwarz, 1979); computer programs (McKeithen, Reitman, Rueter & Hirtle, 1981); basketball (Allard, Graham & Paarsalu, 1980). Some studies, however, have found that experts keep their superiority over novices when the material is memory for randomized sequences of pitch symbols (Sloboda, 1976) and dance sequences (Allard & Starkes, 1991).

A complete lack of difference in memory for random material between experts and non-experts is somewhat counter-intuitive. Simon and Chase (1973) have proposed that 10,000 hours, about ten years, of intense practice and study are necessary to reach a high level of expertise. During their practice and study time experts have undoubtedly met with many situations that are close to “random,”

that is, that contain some infrequently observed features. In the frame of Chase and Simon's (1973b) chunking theory, one could expect that the numerous chunks they have stored in LTM, including some of these unusual features, would allow them to recognize, more often than weak players, familiar chunks that occur adventitiously in random positions, thereby obtaining an advantage in recall. It is also possible that strong players have developed strategies to cope with uncommon situations, which do occur sometimes in their practice. In addition, their familiarity with the materials (for example, in board games, better knowledge of the topology of the board and its attributes) could give them some advantage in comparison with non-experts.

In this paper, we re-examine the recall of random material as a function of level of expertise, emphasizing the chess domain mainly because many empirical data are available there, and because the ELO¹ scale allows players to be ranked much more precisely than just as "experts" and "novices," as is the case in other research domains on expertise. We shall see that there has been some oversimplification of the empirical findings on recall of random chess positions, in the way of ignoring small, but statistically significant, effects.

While random positions have been used mainly as a control of subjects' "general" (non-domain-specific) memory capacity, some researchers have studied them for their own sake. For example, Holding and Reynolds (1982) used semi-random positions to study problem solving in chess and Reynolds (1982) has shown that different degrees of "randomness" may be obtained by manipulating the amount of control that the pieces have over the center.

Some studies have found skill differences in recall performance when the presentation time is sufficiently long. Djakow, Petrowski, and Rudik (1927) presented for one minute a “random” position to Masters and to subjects in a non-chessplaying control group, and found that Masters’ recall was better than control group subjects’. However, two difficulties cloud the interpretation of this study. First, the subjects of the control group did not play chess at all, hence may have been wholly unfamiliar with the material. Second, the position was a chess problem. Chess problems are specially constructed situations where the first goal is to construct esthetic positions and combinations. Although this species of chess is quite different from normal chess games, it employs positions that are far from random. Lories (1987) found an effect of skill with one-minute presentation of the semi-random positions² generated by Holding and Reynolds (1982).

In addition, Saariluoma (1989) used a procedure similar to Chase and Ericsson’s (1982) for the memory of digits, dictating positions at the pace of 2 or 4 seconds per piece. He found that strong players are better in the recall of both game and random positions. Finally, Goldin (1979) and Saariluoma (1984) showed that skilled players perform better than less skilled players in a recognition task, both with game and random positions and both with long presentation times (no limit in the study time in Goldin, 1979) and short presentation times (8 seconds in Saariluoma, 1984).

While it is agreed that Masters do better with random positions in some special memory tasks, like those studied by Saariluoma (1989), the general view among students of expert memory has been that there is no difference in recall with the standard presentation time of five seconds (see for example Cooke, Atlas,

Lane, & Berger, 1993; Ericsson & Charness, 1994; Holding, 1985). However, an analysis of data from various experiments in the literature that used recall of random positions as a control condition makes it clear that strong players show rather reliably a superiority over weaker players. Table 1 lists all the experiments we have found bearing on this question, with the additional criteria that the presentation time should be at most 10 seconds, that the mean number of pieces per position should be at least twenty,³ and that the positions used should be generated by a truly random procedure (see footnote 2).

Insert Table 1 about here

In all cases, except in Chase and Simon's (1973a) study, recall performance increases monotonically as a function of skill. The F -, t - or p -values are not systematically reported in these studies, so we cannot use such meta-analytical approaches as combining tests or computing the overall effect size. We observe, however, that the strongest skill group outperforms the weakest in 12 cases out of 13. Assuming a binomial distribution and equal probability of the strongest group performing better or worse than the weakest, the probability of the strongest group outperforming the weakest in twelve or more cases is .0017. Alternatively, we can analyze the data in Table 1 using a randomized block design, with skill level as independent variable and experiment as blocking variable. Again, skill levels differ reliably [$F(3, 17) = 10.35$, $MS_e = 0.97$, $p < .001$]. Note finally that the type of reconstruction--board and pieces, dictation, or

computer display-- does not influence performance [$F(2, 27) = 0.90$, $MS_e = 2.54$, ns.].

We see that for each experiment, with the exception of Chase and Simon's results (1973a) where the Master performed worse than the novice, the more skilled players do remember more pieces on random boards than the less skilled. That these differences were in most cases not statistically significant may be explained both by the small size of the effect and the small number of subjects in these experiments. In a study (Gobet & Simon, 1995) aimed at extending the present findings, the presentation time was systematically varied from one second to 60 seconds. It was found that skill differences were present at all presentation times, although they were larger with long presentation times.

We must deal at the outset with a plausible explanation for this skill difference, suggested to us by Neil Charness: that chess masters guess better and more than novices. That strong players guess better with random material is not likely, as Jongman (1968) found that strong players were not better than weaker players in guessing the zero-order probability of piece location in game positions. Whether strong players guess more can be checked against the number of errors of commission they make. As data on errors of commission are not available for all the studies listed in Table 1, we will restrict our analysis to experiments run by our research group. In Gobet and Simon (in press-b), Masters, Experts, and Class A players committed, on average 2.4, 5.5 and 3.9 errors of commission with random positions presented for 5 second each. The differences were not significant statistically. In Gobet and Simon (1995), where the presentation time was varied from 1 sec to 60 sec, players of all skill levels tended to make more

errors of commission with longer presentation times, and Masters tended to make fewer errors of commission than Experts and Class A players. Again, the differences were not statistically significant. With five seconds, the respective errors of commission were 1.0, 3.1, 3.4, for Masters, Experts, and Class A players, respectively. Based on these data and on Jongman's results, we can safely reject the hypothesis that the differences we found were due to stronger players guessing more or better than weaker players.

Thus, while randomization does severely reduce the recall of strong players, it does not destroy completely their superiority over weaker players. Note that this skill difference in recall with random positions is small (roughly one piece per additional 400 ELO points), much less than for the recall of game positions, where an increase of 400 ELO points yields typically an increase of about five pieces. Figure 1, which summarize the data of Table 1 and gives the mean number of pieces correct for the game positions obtained in the same experiments, makes it clear that the slopes are different in the two cases.

Insert Figure 1 about here

Does this finding impair the prevailing theory that experts' superiority of recall (especially with game positions) rests on their superior store of chess knowledge? We conclude that it does not. First, the few studies that have investigated the general cognitive abilities of chessplayers outside the domain of

chess (see Cranberg and Albert, 1988, De Groot, [1946] 1978, or Gobet, 1993, for reviews) have found no superiority for chessplayers in comparison with non-chessplayers, nor is there a correlation between level of skill and general cognitive abilities. For example, Djakow et al. (1927), subjected top-level masters to a battery of psychological tests, including a visual memory test using 8x8 matrices having or not having a resemblance to chess boards. Masters performed better than control subjects when tests involved chess material, but not in tests unrelated to chess. (A later exception is the subject VP studied by Hunt and Love, 1972, who was both a mnemonist and a strong chessplayer.)

Second, the skill difference with random positions is predicted by Chase and Simon's (1973b) chunking theory, because Masters, having a larger repertoire of chunks in LTM, are more likely to find rare patterns in LTM than weaker players.⁴ This is confirmed by computer simulations (Gobet and Simon, in press-b), where the program's recall of random positions increased logarithmically as a function of the number of chunks that had been learned through the study of game positions.

What processes allow strong players to perform better than weak players with random positions? We have suggested three possibilities: (1) a large database of chunks in LTM, occasionally allowing the recognition of stored patterns that occur by chance in random positions; (2) the possession of strategies for coping with uncommon positions; (3) better knowledge of the topology of the chessboard. We have just seen that the first hypothesis is supported by computer simulation. We are aware of no specific evidence for the second hypothesis. The

third hypothesis, of better knowledge of the topology of the board, is supported by Saariluoma's (1991) data, which show that Masters are better than novices at deciding whether a square, denoted by its algebraic notation (e.g. "e4") is White or Black. Following this hypothesis, Masters possess chunks of squares (with or without pieces on them), that may be used in random positions to organize patterns of pieces. For example, Masters may know that the squares "a1" and "c2" are at a Knight's distance, and may use this schema of squares to encode the pattern "White Pawn a1, Black Pawn c2," even though, because the rules of the game prohibit Pawns on the first rank, they are not likely to have learned this precise chunk through past experience.

In this paper, we have reviewed several chess recall studies that used random positions as a control task, with a presentation time ranging from three to ten seconds. We found that, contrary to the general opinion, the recall of random position varies somewhat as a function of chess skill; and we proposed an explanation of this effect that is supported by our computer simulations. Although the absolute difference between skill levels is small, its presence rules out some theoretical accounts of chess expertise that have been proposed as alternatives to the chunking theory. Without an ad hoc assumption, perhaps familiarity with chess material, it is hard to see how theories of chess skill based mainly on level of processing (Lane & Robertson, 1979) or high-level knowledge (Cooke et al., 1993) can account for this observed, if modest, effect with random positions. By accessing chunks (not information at a "deeper level") already stored

in LTM, our computer model of the expert occasionally detects and recalls adventitious patterns on the random boards. The level-of-processing theory would require accessing "higher knowledge" in LTM, but it is not obvious what "higher knowledge" could be used in recalling random positions, which, by construction, have no deeper structure.

As with chess, most studies on experts' memory in other tasks have used small numbers of subjects (Charness, 1988; Gobet, 1993). It is a question for future research whether the lack of difference between experts and non-experts found in various domains of expertise when random material was used is genuine or is due to the low power of the experimental design used in these studies.

Taken together with results on the memory for sequence of pitch symbols (Sloboda, 1976) and for dance sequences (Allard & Starkes, 1991), the studies reviewed in this paper indicate that the relation between expertise and "meaningless" material taken from the domain of expertise may be somewhat more complex than has been thought previously, but that the notion of chunking is able to account for most of the phenomena, quantitatively as well as qualitatively.

References

- Allard, F., Graham, S. & Paarsalu, M. E. (1980). Perception in sport: Basketball. *Journal of Sport Psychology*, 2, 22-30.
- Allard, F., & Starkes, J. L. (1991). Motor-skill experts in sports, dance and other domains. In K. A. Ericsson (Eds.), & J. Smith, *Toward a general theory of expertise. Prospects and limits*. Cambridge: Cambridge University Press.
- Anderson, J. R. (1990). *Cognitive psychology and its implications*. New York: Freeman.
- Charness, N. (1979). Components of skill in bridge. *Canadian Journal of Psychology*, 33, 1-16.
- Charness, N. (1988). Expertise in chess, music and physics : A cognitive perspective. In Obler, L. K. & Fein D. A. (Eds.) *The exceptional brain : The neuropsychology of talent and special abilities*. New-York, Guilford Press.
- Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 16). New York: Academic Press.
- Chase, W. G., & Simon, H. A. (1973a). Perception in chess. *Cognitive Psychology*, 4, 55-81.
- Chase, W. G., & Simon, H. A. (1973b). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press.

- Cooke, N., J., Atlas, R. S., Lane, D. M., & Berger, R. C. (1993). Role of high-level knowledge in memory for chess positions. *American Journal of Psychology, 106*, 321-351.
- Cranberg, L. D., & Albert, M. L. (1988). The chess mind. In L. K. Obler & D. Fein (Eds.), *The exceptional brain. Neuropsychology of talent and special abilities*. New York: Guilford.
- De Groot, A. D. (1946). *Het denken van den schaker*. Amsterdam, Noord Hollandsche.
- De Groot, A. D. (1978). *Thought and choice in chess* (2nd Ed.). The Hague: Mouton Publishers.
- Djakow, I. N., Petrowski, N. W., & Rudik, P. A. (1927). *Psychologie des Schachspiels*. Berlin: de Gruyter.
- Egan, D. E., & Schwartz, E. J. (1979). Chunking in recall of symbolic drawings. *Memory & Cognition, 7*, 149-158.
- Engle, R. W., & Bukstel, L. (1978). Memory processes among bridge players of differing expertise. *American Journal of Psychology, 91*, 673-689.
- Ericsson, K. A., & Charness, N. (1994). Expert performance. Its structure and acquisition. *American Psychologist, 49*, 725-747.
- Frey, P. W., & Adesman, P. (1976). Recall memory for visually presented chess positions. *Memory and Cognition, 4*, 541-547.
- Gobet, F. (1993). *Les mémoires d'un joueur d'échecs*. Fribourg (Switzerland): Editions Universitaires.

- Gobet, F. & Simon, H. A. (1995). *Role of presentation time in recall of game and random chess positions*. (Complex Information Paper #524). Pittsburgh: Carnegie Mellon University, Department of Psychology.
- Gobet, F. & Simon, H. A. (in press-a). Templates in chess memory: a mechanism for recalling several boards. *Cognitive Psychology*.
- Gobet, F. & Simon, H. A. (in press-b). Recall of random and distorted chess positions: Implications for the theory of expertise. *Memory & Cognition*.
- Gold, A. & Opwis, K. (1992). Methoden zur empirischen Analyse von Chunks beim Reproduzieren von Schachstellungen. *Sprache & Cognition, 11*, 1-13.
- Goldin, S. E. (1979). Recognition memory for chess positions: Some preliminary research. *American Journal of Psychology, 92*, 19-31.
- Holding, D. H. (1985). *The psychology of chess skill*. Hillsdale, NJ: Erlbaum.
- Holding, D., & Reynolds, R. (1982). Recall or evaluation of chess positions as determinants of chess skill. *Memory & Cognition, 10*, 237-242.
- Hunt, E., & Love, T. (1972). How good can memory be? In A. W. Melton, & E. Martin (Eds.), *Coding processes in human memory*. New York: John Wiley..
- Jongman, R. W. (1968). *Het oog van de meester*. Assen: Van Gorcum.
- Lane, D. M., & Robertson, L. (1979). The generality of the levels of processing hypothesis: An application to memory for chess positions. *Memory & Cognition, 7*, 253-256.
- Lesgold, A. (1988). Problem Solving. In R. J. Sternberg, & E. E. Smith, *The psychology of human thought*. Cambridge: Cambridge University Press.

- Lories, G. (1987). Recall of random and non random chess positions in strong and weak chess players. *Psychologica Belgica*, 27, 153-159.
- McKeithen, K. B., Reitman, J. S., Rueter, H. H., & Hirtle, S. C. (1981). Knowledge organization and skill differences in computer programmer. *Cognitive Psychology*, 13, 307-325.
- Reitman, J. S. (1976). Skilled perception in go: deducing memory structures from inter-response times. *Cognitive Psychology*, 8, 336-356.
- Reynolds, R. I. (1982). Search heuristics of chess players of different calibers. *American Journal of Psychology*, 95, 383-392.
- Saariluoma, P. (1984). *Coding problem spaces in chess: A psychological study*. Commentationes scientiarum socialium 23. Turku: Societas Scientiarum Fennica.
- Saariluoma, P. (1985). Chess players' intake of task-relevant cues. *Memory & Cognition*, 13, 385-391.
- Saariluoma, P. (1989). Chess players' recall of auditorily presented chess positions. *European Journal of Cognitive Psychology*, 1, 309-320.
- Saariluoma, P. (1991). Aspects of skilled imagery in blindfold chess. *Acta Psychologica*, 77, 65-89.
- Saariluoma, P. (1994). Location coding in chess. *The Quarterly Journal of Experimental Psychology*, 47A, 607-630.
- Simon, H. A., & Chase, W.G. (1973). Skill in chess, *American Scientist*, 61, 393-403.

- Sloboda, J. A. (1976). Visual perception of musical notation: Registering pitch symbols in memory. *Quarterly Journal of Experimental Psychology*, 28, 1-16.
- Vicente, K. J. & De Groot, A. D. (1990). The memory recall paradigm: straightening out the historical record, *American Psychologist*, February, 285-287.
- Wolff, A. S., Mitchell, D. H. & Frey, P. W. (1984). Perceptual skill in the game of Othello. *Journal of Psychology*, 118, 7-16.

Authors note

Preparation of this article was supported by grant no 8210-30606 from the Swiss National Funds of Scientific Research to the first author and grant no DBS-912-1027 from the National Science Foundation to the second author. Correspondence concerning this article should be addressed to Herbert A. Simon, Department of Psychology, Carnegie Mellon University, Pittsburgh, Pennsylvania, 15213.

The authors extend their thanks to Neil Charness, David Lane, Howard Richman, Pertti Saariluoma, Jim Staszewski, and Shmuel Ur for valuable comments on parts of this research.

Footnotes

¹Chess players are ranked according to the ELO rating, an interval scale used internationally. Grandmasters are usually rated above 2500 ELO, International Masters above 2400, Masters between 2200 and 2400, Experts between 2000 and 2200, Class A players between 1800 and 2000, Class B players between 1600 and 1800, and so on. The ratings of the US Chess Federation (USCF) are slightly higher than international ELO ratings. Given the large interval classes we will use in this paper, we may ignore the latter difference.

²Lories' study is difficult to interpret because the positions he used, taken from Holding and Reynolds (1982), are not really random. Some (semantic) constraints were applied in generating the positions, such as no Pawn on the first or eighth rank or no piece attacked without being defended. As a matter of fact, a statistical analysis shows that equiprobability of White and Black pieces' distribution on the board may be rejected at $p < .001$ (Gobet, 1993).

³When the range was given instead of the mean number of pieces per position in the description of the experiment, we took the range midpoint.

⁴Though the patterns are produced by chance, they will be recognized if already stored in memory, analogous to noticing thirteen spades in a hand dealt at bridge.

Figure Caption

Figure 1. Mean number of pieces placed correctly as a function of type of positions (game or random) and skill level.

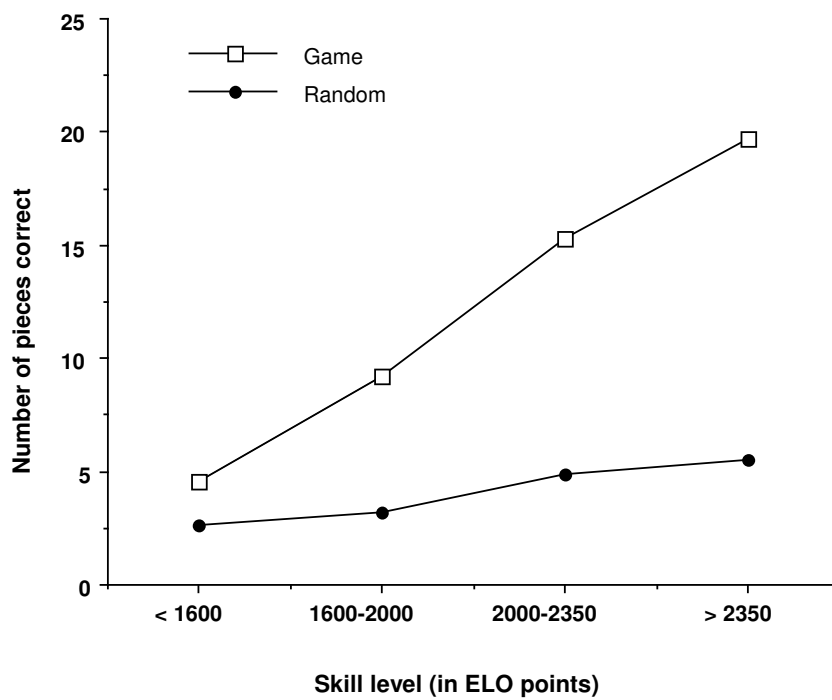


Table 1

Number of pieces correctly replaced for random positions as a function of skill in thirteen experiments. For each experiment, the table lists, in order: the source, the number of subjects, the presentation time in seconds, the reconstruction mode, and the number of pieces correctly replaced.

Source	N of subjects	Presen- tation Time (in seconds)	Recons- truction Mode	Mean Rating (in Elo points) ^a			
				<1600	1600-2000	2000-2350	>2350
bh Chase & Simon (1973a)	3	5	board	3.0	2.5		2.0
bh Frey & Adesman (1976)	13	8	board	2.0	2.5		
di Saariluoma (1984), exp. 3	4	5	board	2.4	3.8		

Skilled chessplayers and random positions

21

d	Saariluoma (1984), exp. 4	4	5	board	2.5		5.0	7.3
dh	Saariluoma (1985), exp. 3	9	5	?	2.6	3.8	4.6	
ech	Gold & Opwis (1992)	40	10	board	3.0	4.3		
dfh	Saariluoma (1994), exp. 1	12	5	board	3.0		5.8	
edh	Saariluoma (1994), exp. 2	9	3	verbal	2.4	3.4	4.8	
dh	Saariluoma (1994), exp. 3	8	5	verbal	2.4		6.0	
dgh	Saariluoma (1994), exp. 4	10	3	verbal	2.4	2.9		
cg	Gobet & Simon (1995), exp. 1	21	5	computer		3.4	5.2	8.4
c	Gobet & Simon (in press-a), exp. 1	13	5	computer		3.0	4.0	5.2
c	Gobet & Simon (in press-b), exp. 2	25	5	computer		3.1	3.4	3.6

(End of Table 1)

^a The group mean rating is used for classification

^b USCF rating is used

^c International rating, or equivalent, is used

^d Finnish rating is used

^e The difference between skill levels is significant at the .05 level.

^f This corrects an error in Saariluoma's (1994) figure 2, where the labels for the TT and random positions have been swapped (Saariluoma, 1995, personal communication).

^g The experiment used other presentation times as well.

^h Values estimated from graph

ⁱ Experimental conditions with 20 and 25 pieces pooled.