

P.C.R. Lane and F. Gobet,

'Perception in chess and beyond: Commentary on Linhares and Freitas (2010)',

New Ideas in Psychology, 29:156-61, 2011

Abstract

Linhares and Freitas (2010; LF) argue that experts use analogical or semantic similarity, similarities that are not available from direct surface representations. LF make their case using a critique of Chase and Simon (1973b) and the presentation of a few chess positions and examples from other domains. Their conclusion is that models such as CHREST (Gobet et al., 2001) and theories such as the chunking theory (Chase & Simon, 1973b) and the template theory (Gobet & Simon, 1996) are inadequate for dealing with these issues. They propose an alternative paradigm, which they call “experience recognition.” Although we find this issue an interesting one, the separation between pattern *recognition* and problem *solving* is a lot more complex than LF portray. We instead suggest that a “revolution” in our to date successful modelling is not necessary. Especially in the chess domain, LF’s examples do not make the point they claim. Furthermore, their criticisms of CS are incorrect, and they have failed to mention a large number of experimental results that have supported the hypothesis of location-specific encodings. Although we agree that experts use semantic information and similarities, these ideas already possess analogues in CHREST, which can form the basis of further evolution of the theory.

Keywords

Analogy; Chunking; Cognitive models; Cognitive psychology; Memory; Pattern Recognition; Perception, Template

Perception in Chess and Beyond: Commentary on Linhares and Freitas (2010)

Ever since the seminal work of De Groot (1978/1946) on chess, the link between perception and abstraction has been a central question in psychology. First, there is the question of what is at the heart of (expert) cognition: perception based on surface similarity vs. perception based on analogy. Second, there is the question of whether perception relies on a large number of long-term memory structures acquired over a long period of time (typically more than 10 years with experts), as proposed by chunk-based theories (Chase & Simon, 1973b; Gobet & Simon, 1996b), or whether perception relies on the rapid creation of new chunks, for example using a small number of combinations of fluid abstract roles, as proposed by Linhares (2005) and Linhares and Brum (2007).

Linhares and Freitas (2010; hereafter LF) analyse the classic paper by Chase and Simon (1973b; hereafter CS) and later modelling work using CHREST (De Groot & Gobet, 1996; Gobet et al., 2001; Gobet & Waters, 2003) and CHUMP (Gobet & Jansen, 1994), and discuss research on the hypothesis that analogy lies at the core of cognition. They conclude that undue attention has been given to the idea of pattern recognition, and that the field should instead focus on “experience recognition.” More specifically, they argue that Chase and Simon’s analysis of chunks is flawed, and that chunking theory, CHREST and CHUMP and other theories based on pattern recognition are invalid because they focus on surface information (the assumption of location coding) and do not consider abstract and semantic features. In this commentary, we show that LF’s criticisms are deeply flawed.

The Assumption of Location Coding

We begin with what LF see as “perhaps the most profound problem” (p. 69) with CHREST and related models: their lack of flexibility, and in particular their assumption that knowledge is stored with information about the specific location. When making their case, LF often use

illustrative chess positions and ask the reader to rely on their intuitions. With respect to the issue of location coding, they invite the reader to imagine that all pieces in their position 10 (reproduced here in Figure 1a) are shifted horizontally or vertically (or both). Then, they write, (LF, p. 69): “But is this shifted position actually different from the original in any significant sense? No: experts have reported to us that this shifting of position 10 ‘means nothing, it’s the very same position’ (Linhares & Brum, 2007). Every piece may have moved, but in essence, the position is the same – and so is the strategy for play.” One of us is an international chess master and, like any master worth their salt, cringes when reading this. Consider the positions (b) and (c) in Figure 1. In position (b), the pieces have been shifted down and to the left; the black King is stalemated (no legal move is possible for Black), and the outcome of the game is a draw rather than a win as in position (a). In position (c), the pieces have been shifted down and to the right (except for the white King, which cannot be shifted to the right as it would be outside of the bounds of the board). Actually, the position of the white King is critical. If it is located on h8 or h7, the position is won for White. If it is located anywhere else the position is a draw, because Black keeps his King on e7 and f8, and the only plan for winning is to bring the King to the squares d8, d7, or d6 (there is no path to h8 or h7), at which point the black King is stalemate.

Insert Figure 1 about here

Insert Figure 2 about here

The reader might object that, in this case, LF’s statement holds in most shiftings of the pieces. This is true, but consider now their position 20, reproduced here as Figure 2a. The key pattern here is (kb8, rc8, pa7, pb7, Nc7, Qf4), and the solution is 1.Na6 double-check ka8 2.Qb8 check rxb8 3.Nc7 checkmate. Now shift this pattern anywhere on the board, for example as in Figure 2b, and there is no checkmate. (If you allow for mirror-image symmetry and change of colour, this combination is possible only in the other three corners of the board.) This combination

relies on the mobility of the black King being limited by the side of the board: there is no ninth row and zeroth column. Indeed, position 2b is winning for Black while position 2a was winning for White.

The assumption of location encoding has been challenged before LF by Holding (1985), who suggested that the number of chunks proposed by Chase and Simon was exaggerated, and could be reduced by using symmetries on the board. This suggestion has not been supported by experiments. Saariluoma (1994) showed that swapping quadrants of positions impaired recall *within* the swapped quadrants; Gobet and Simon (1996a) showed that mirror images impaired recall; and Didierjean, Cauzinille-Marmèche and Savina (1999) found that players could not generalise the smothered checkmate pattern (as in Figure 2a) to its mirror image.

In the examples above, shifting the complete chess position leads to radical changes in its evaluation. This is because, with a complete position, the edges of the board play a significant role. In contrast, problem solving in the small may require a different treatment: for example, learning that a white pawn on d4 may take a black knight on e5 could be generalised to a position with a white pawn on e4 and a black knight on f5. We agree that in *local problem solving* cases such as this, the location-encoding representation removes the possibility of a natural generalisation.

The distinction between global and local problem solving means that LF's objections do not identify any flaw in the CHREST architecture, for several reasons. First, the discrimination network within CHREST is not the complete repository of the model's chess knowledge. The eye heuristics also rely on domain expertise, and suggest fixations based on local information, such as attacking/defending moves and natural groupings or salient pieces. Second, CHREST to date has not been intended as a model of chess *playing* but only of chess *perception* and *memory*. The focus on chess memory, especially of expert players, has led us to develop a model capable of dealing with the complete chessboard. Small refinements to the internal representations can be used to provide a bridge between local relations and global location-encodings; this bridge would enable some local generalisations to be made fairly easily, but this is not the place to discuss alternative

representations.

Corrections of Critique of CS

LF state their view of CS very clearly (p. 72, italics in original): “[LF] is not making a minor claim about some obscure technical error: *all the results* in the 20 pages following page 61 are claimed to be invalid.” These 20 pages consist of the presentation of a methodology for identifying chunks using both a copy and a recall task, and the presentation and analysis of the results they obtained.¹ LF’s key claim is that, in CS, “performance on the tasks is indistinguishable between a master and a beginner” (p. 72). LF’s argument is thus that there is *no* difference at all in performance between the master and the beginner with respect to all analyses related to chunking. However, a cursory look at CS shows that this claim is plainly wrong.

Latencies. With respect to the analysis of within and between glance latencies, LF write that “The results were unequivocal: the data *was exactly the same for masters and beginners* (see figs. 3 and 4 of that paper). They pointed this out clearly: [Perception task, p. 65] ‘The first thing to notice is that the data are quite similar for all subjects. ...’” (LF, p. 71, italics in original).

This claim is simply incorrect. On p. 62, CS explicitly discuss the skill difference for the perception task (lower part of their Figure 3): “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 [sic] significant ($p < .05$) when tested against a pooled error term.” The passage that LF quote (starting with “The first thing to notice...”) from CS is literally out of context: the passage refers to the discussion of Tables 1 and 2 (which deal with the probabilities and latencies as a function of the patterns of relations – attack, defense, etc.), and not, as implied by LF, to Figures 3 and 4.

Contents of chunks. As made clear by the previous quotation from CS, the pattern of latencies and probabilities as a function of the sixteen possible combinations of attack, defense, similarity and color, did not differentiate between the three skill levels. The two tables, however,

show clear differences in latencies and probabilities depending on whether the pairs of pieces were within or between glance. A similar pattern was found with the placements from the memory task, now categorized as a function of whether they were placed with an interval of less or more than two seconds. This was an essential result for the claim that their method was able to identify chunk boundaries. But, again, there was no difference between the three players, suggesting that they used the same perceptual building blocks for constructing chunks (more about this later.)

Size of chunk. CS found that chunk size differed in the memory task between the three players (2.5, 2.1, and 1.9 pieces, for the master, Class A player, and beginner, respectively) and we agree with LF and others (e.g., Holding, 1985) that the master's chunks were small. CS (p. 76) note 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." While finding different sizes in the two tasks would have been more convincing, this statement is plausible, given the statistically reliable difference found with the between-glance latencies (see above). Thus, LF's criticism of this aspect of the results is unfounded.

Number of chunks. Finally, LF incorrectly claim that performance between the master and the beginner cannot be distinguished with respect to the number of chunks. Although the number of chunks was well within the postulated size of short-term memory (7 ± 2 chunks), it differed between the three participants (7.7, 5.7 and 5.3 pieces for the master, class A player and beginner, respectively), and Chase and Simon considered this result as a serious challenge for their theory.

The content of chunks warrants further comments. Referring to the within- and between-chunk probabilities of the different chess relations in the copy and recall task, CS (p. 68) note that "these probabilities are informative about the underlying structures that the subjects are perceiving." LF take issue with this conclusion (p. 72): "The probabilities cannot be informative if there are no differences between those that have the "underlying structures" (i.e., chunks) and those that lack them." Here, LF refer to chunks as long-term memory structures, as proposed by the theory, not as the chunks inferred by CS's methodology, which might also include purely perceptual chunks, as

this must be the case with the beginner. In CS, these data were informative, because they showed that the same low-level perceptual features characterize the chunks of players of different levels. As noted by CS, “subjects are noticing the pawn structure, clusters of pieces of the same color, and attack and defense relations over small spatial distances” (p. 68). Interestingly, relations of attack, in particular over long distances, are rare.² These data are key to reach the conclusion that chunks are “perceptual chunks.”

Later Evidence for Chunking

CS’s experiment generated considerable research, which LF have not mentioned, on the links between perception and cognition (for details, see Gobet, de Voogt, & Retschitzki, 2004; Gobet & Simon, 1998a, 1998b); this further research was part of the development of CHREST (which did not depend solely on CS, as LF imply). Two replications of the CS experiment (Gobet & Clarkson, 2004; Gobet & Simon, 1998a), with larger samples than the original study, added support to the result that the pattern of relations is different between and within chunks, and to the two-second boundary for identifying chunks. These replications also showed that the two main anomalies mentioned above disappear when the experiment is carried out with a computer display rather than with physical board and pieces. That is, the replications show that masters use large chunks (up to 15 pieces in Gobet & Clarkson, 2004) and that the number of chunks is similar between skill levels in the memory task (less than three chunks – a smaller estimate than that made by CS). A likely explanation for these discrepancies is that, with CS methodology, the size of the chunks is limited by the number of pieces that the hand can hold, while this is not the case with a computer display.

In two studies (Charness, 1976; Frey & Adelman, 1976), the presence of chunks was tested experimentally. Pieces were presented either grouped using CS chunking relations, ordered by columns, or dictated in a random order. The chunk presentation obtained the best recall, as predicted by the theory.

One could criticize the fact that much research has been carried out on chess memory, while

the essence of chess playing is finding good moves, not memorising boards. However, memorising boards is far from lacking ecological validity, as chess masters often discuss other players' games in progress, and typically do so using their memory of the board position. In addition, Chase and Simon themselves (1973a) carried out a number of other tasks, such as memory for sequences of moves and memory for games – both of which are essential activities for becoming a chess expert. Similarly, and starting with De Groot's (1946) study, chess research has also focused on problem solving (for recent examples, see Bilalić, Mcleod, & Gobet, 2008; Jeremic, Vukmirovic, & Radojicic, 2010). Finally, the role of chunking and pattern recognition has been established in many other domains of expertise, including games, sports, and science (for reviews, see Ericsson et al., 2006). Indeed, one of the strengths of Chase and Simon's theory, and other theories based on chunking, is to have shown that the same mechanisms not only account for empirical data from memory experiments but also explain how experts find meaning in problem situations and reach good solutions while being highly selective in their search.

Corrections of Critique of CHREST/CHUMP

LF have misunderstood a few key points about CHREST and CHUMP, which it is important to correct here. The first area is the relation of CHREST to semantic knowledge, such as that a position has the theme of king-opposition. Classic theories of chunking (of which CHREST is a development) have always been strong on the need to include meaning, analogy and abstract concepts (see, for example, De Groot & Gobet, 1996; Didierjean et al., 1999; Freyhoff, Gruber, & Ziegler, 1992; Gobet et al., 2001). The template theory, which CHREST implements, directly addresses the issue of how high-level knowledge can be accessed when confronted with a specific chess position (Gobet & Simon, 1996b); templates account for the generality of themes in both a specific and an abstract sense. This area has been developed with CHREST, but not so much with chess (see Gobet & Lane, 2005; Lane, Sykes, & Gobet, 2003).

The second area regards CHUMP, and the ability of CHREST to be a model of problem-solving. LF present an extended analysis on the unlikely ability of CHUMP – a simple program that

chooses moves using pattern recognition only – to locate a good move in a complex situation. LF have both promoted CHUMP to the level of a coherent model of expert chess playing, and also misunderstood its basic operation. CHUMP is not the model of move selection which the current authors envisage as the natural association with CHREST; the appropriate model is SEARCH (Gobet, 1997). SEARCH is a probabilistic and abstract model that iteratively looks ahead from the current position using pattern recognition to retrieve chunks that are used to generate a move. SEARCH will consider several such episodes of analysis, as time and resources permit.

To correct a specific concern of LF: CHREST does not need to acquire a unique chunk or template for every position – instead, a set of chunks or templates will be retrieved for any position and a candidate move selected from a range of sources.

Conclusion: Evolution or Revolution?

Although you can expect the present authors to have a bias towards their own theory, we feel there is enough existing evidence to support the continuing evolution of CHREST towards a more complete theory of human perception and problem solving. At the least, we do not believe LF have presented enough evidence to reject the conclusions of CS or the development of chunking theory and related models of expertise. In addition, it is unlikely that one single mechanism – analogy – explains an activity as complex as chess skill, and more likely that a number of mechanisms, such as those implemented in CHREST and SEARCH (including implicit learning, pattern recognition, visual search, and look-ahead search) are at play. We suggest that the fundamental result of de Groot, that expertise is driven by memory of items in a domain, will remain a cornerstone of expertise research. The challenge, currently taken up by CHREST, is to see how this memory can be acquired and built up into the more abstract patterns of thinking and comparison referred to above. We believe the only way to meet that challenge is through evolving our current understanding of human memory into a more powerful theory, not by discarding what has gone before in a drastic revolution.

References

- Bilalić, M., Mcleod, P., & Gobet, F. (2008). Inflexibility of experts - Reality or myth? Quantifying the Einstellung effect in chess masters. *Cognitive Psychology*, *56*, 73-102.
- Charness, N. (1976). Memory for chess positions: Resistance to interference. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 641-653.
- Chase, W. G., & Simon, H. A. (1973a). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215-281). New York: Academic Press.
- Chase, W. G., & Simon, H. A. (1973b). Perception in chess. *Cognitive Psychology*, *4*, 55-81.
- De Groot, A. D. (1978). *Thought and choice in chess (first Dutch edition in 1946)*. The Hague: Mouton Publishers.
- De Groot, A. D., & Gobet, F. (1996). *Perception and memory in chess*. Assen: Van Gorcum.
- Didierjean, A., Cauzinille-Marmèche, E., & Savina, Y. (1999). Learning from examples: Case-based reasoning in chess for novices. *Cahiers de Psychologie Cognitive-Current Psychology of Cognition*, *18*, 337-361.
- Ericsson, K. A., Charness, N., Feltovich, P. J., & Hoffman, R. R. (2006). *The Cambridge handbook of expertise and expert performance*. New York, NY: Cambridge University Press.
- Frey, P. W., & Adesman, P. (1976). Recall memory for visually presented chess positions. *Memory and Cognition*, *4*, 541-547.
- Freyhoff, H., Gruber, H., & Ziegler, A. (1992). Expertise and hierarchical knowledge representation in chess. *Psychological Research*, *54*, 32-37.
- Gobet, F. (1997). A pattern-recognition theory of search in expert problem solving. *Thinking and Reasoning*, *3*, 291-313.
- Gobet, F., & Clarkson, G. (2004). Chunks in expert memory: Evidence for the magical number four... or is it two? *Memory*, *12*, 732-747.
- Gobet, F., de Voogt, A. J., & Retschitzki, J. (2004). *Moves in mind: The psychology of board games*. Hove, UK: Psychology Press.

- Gobet, F., & Jansen, P. J. (1994). Towards a chess program based on a model of human memory. In H. J. van den Herik, I. S. Herschberg & J. W. H. M. Uiterwijk (Eds.), *Advances in Computer Chess 7* (pp. 35-60). Maastricht: University of Limburg Press.
- Gobet, F., & Lane, P. C. R. (2005). The CHREST architecture of cognition: Listening to empirical data. In D. Davis (Ed.), *Visions of mind: Architectures for cognition and affect* (pp. 204-224). Hershey, PA: IPS.
- Gobet, F., Lane, P. C. R., Croker, S., Cheng, P. C. H., Jones, G., Oliver, I., et al. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences*, 5, 236-243.
- Gobet, F., & Simon, H. A. (1996a). Recall of random and distorted positions. Implications for the theory of expertise. *Memory & Cognition*, 24, 493-503.
- Gobet, F., & Simon, H. A. (1996b). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology*, 31, 1-40.
- Gobet, F., & Simon, H. A. (1998a). Expert chess memory: Revisiting the chunking hypothesis. *Memory*, 6, 225-255.
- Gobet, F., & Simon, H. A. (1998b). Pattern recognition makes search possible: Comments on Holding (1992). *Psychological Research*, 61, 204-208.
- Gobet, F., & Waters, A. J. (2003). The role of constraints in expert memory. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 29, 1082-1094.
- Holding, D. H. (1985). *The psychology of chess skill*. Hillsdale, NJ: Erlbaum.
- Jeremic, V., Vukmirovic, D., & Radojicic, Z. (2010). Does playing blindfold chess reduce the quality of game: Comments on Chabris and Hearst (2003). *Cognitive Science*, 34, 1-9.
- Lane, P. C. R., Sykes, A. K., & Gobet, F. (2003). Combining low-level perception with expectations in CHREST. In F. Schmalhofer, R. M. Young & G. Katz (Eds.), *Proceedings of EuroCogSci 03: The European Cognitive Science Conference 2003* (pp. 205-210). Mahwah, NJ: Erlbaum.
- Linhares, A. (2005). An active symbols theory of chess intuition. *Minds and Machines*, 15, 131-181.

- Linhares, A., & Brum, P. (2007). Understanding our understanding of strategic scenarios: What role do chunks play? *Cognitive Science*, *31*, 989-1007.
- Linhares, A., & Freitas, A. E. T. A. (2010). Questioning Chase and Simon's (1973) "Perception in chess": The "experience recognition" hypothesis. *New Ideas in Psychology*, *28*, 64-78.
- Saariluoma, P. (1994). Location coding in chess. *The Quarterly Journal of Experimental Psychology*, *47A*, 607-630.

Footnotes

¹ LF's description of CS experiment is not particularly clear. For example, the reader might have been confused by the five labels used for the three players used by CS: master, Class A player, Class B player, Class C player, and beginner. In particular, the beginner is sometimes called beginner, sometimes Class B player, and sometimes Class C player. The real skill levels were master, Class A player, and beginner. Incidentally, some of the passages from CS are not correctly quoted by LF, as CS used the labels "A" (for class A player) and "B" (for beginner) rather than "class A" or "class B."

² It might be worth mentioning that, in a task consisting of memorising sequences of moves, relations of defence and attack are used much more often than in the memory for static positions (Chase & Simon, 1973a).

Figures

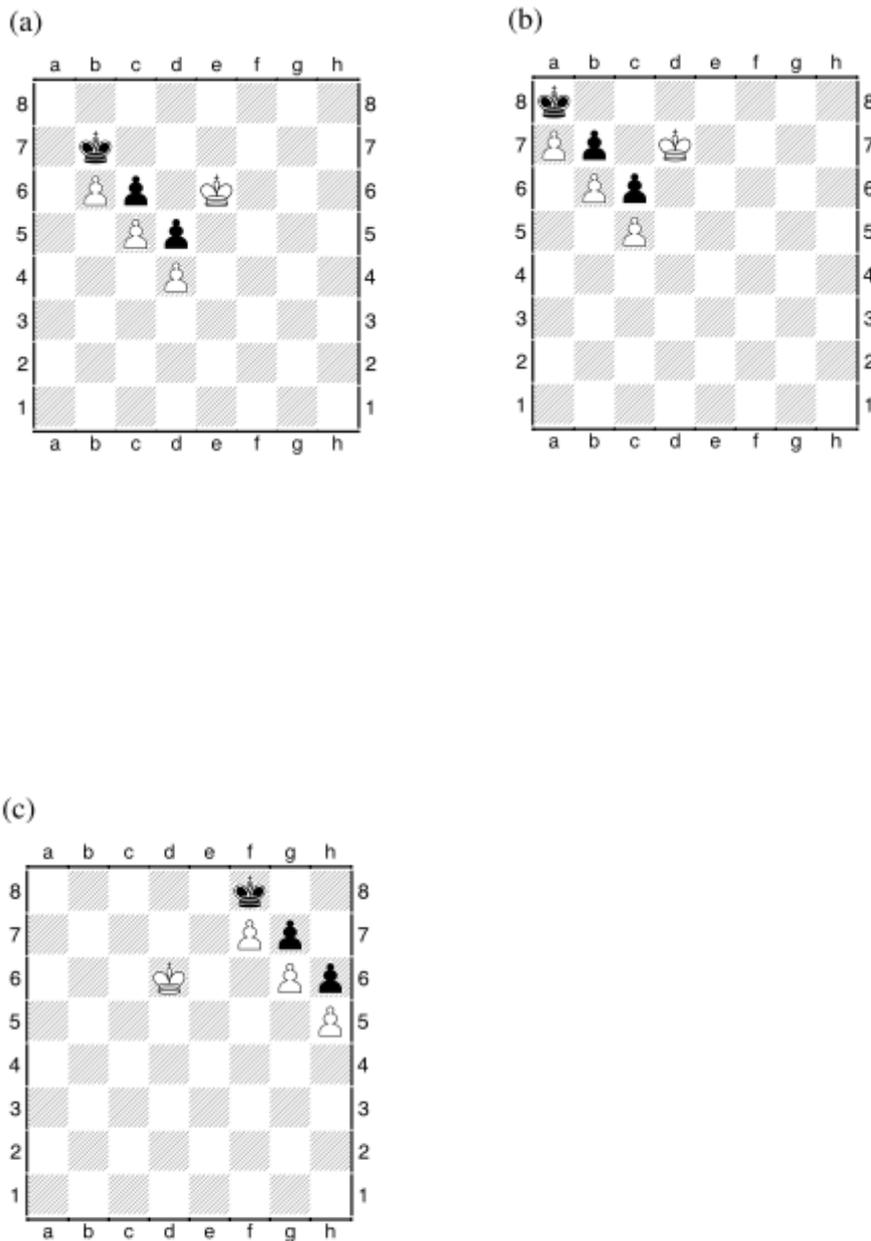
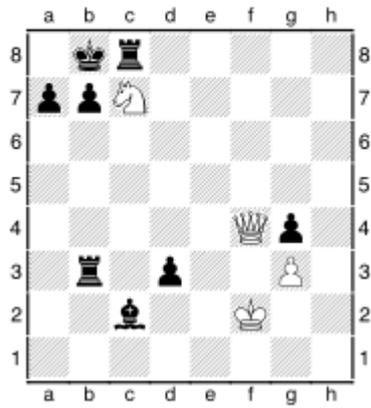


Figure 1. Example of the importance of absolute location in chess. Position (a), taken from Linhares and Freitas (2010), is a win for White, but White cannot win in the shifted positions (b) and (c) (see text). Readers who do not play chess might consider the following example: in basketball, a configuration of two players of team A attacking a player of team B holding the ball has a different meaning depending whether these players are under the basket of team A or in the centre of the field.

(a)



(b)

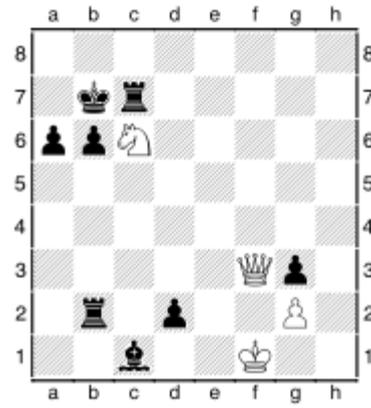


Figure 2. Another example of the importance of absolute location in chess. Position (a), taken from Linhares and Freitas (2010), is a win for White (assuming White plays first), but Black is winning in the shifted position (b). See text for detail.