# Checking Chess Checks with Chunks: A Model of Simple Check Detection

**Richard Ll. Smith (rls@hwyl.org)**
School of Social Sciences, Brunel University,
Uxbridge, UB8 3PH UK

**Fernand Gobet (fernand.gobet@brunel.ac.uk)**
School of Social Sciences, Brunel University,
Uxbridge, UB8 3PH UK

**Peter C.R. Lane (peter.lane@bcs.org.uk)**
School of Computer Science, University of Hertfordshire,
Hatfield, AL10 9AB UK

**Abstract**

The procedure by which humans identify checks in check positions is not well understood. We report here our experience in modelling this process with CHREST, a general-purpose cognitive model that has previously successfully captured a variety of attention- and perception-related phenomena. We have attempted to reproduce the results of an experiment investigating the ability of humans to determine checks in simple chess positions. We propose a specific model of how humans perform this experiment, and show that, given certain reasonable assumptions, CHREST can follow this model to create a good reproduction of the data.

**Keywords:** CHREST; cognitive model; cognitive architecture; chess; check detection

## Motivation

In studying the general phenomenon of human perception, we have looked at the specific task of perceiving checks in the game of chess. This task involves a player being presented with a chess position with a requirement to determine whether or not the player's king is being threatened by another piece.

Experiments on human subjects have provided data about how well they perform this task, but we have no good model of how the underlying psychological processes work in this situation. Identifying threats in games is a complex task which explores the process of visual attention when guided by interpretation of higher goals. Understanding these processes may help shed light on a variety of aspects of attention and perception.

Although understanding these processes is a desirable goal in itself, we are also interested in modelling this process as part of a larger project to produce a cognitive model which, whilst operating under human constraints, plays chess in a human-like way. Successfully modelling the check perception process would be a step towards this aim, as well as a verification of the parameters of the model itself.

## Background

Saariluoma conducted a series of experiments (Saariluoma, 1984) relating to the perception abilities of humans through the medium of chess. We are concerned with one experiment in particular: this was to measure how quickly players could determine, given a chess position consisting of a white king and one other black piece, whether or not the king was in check. The subjects of the experiment included chess players with a mix of skill levels: two complete beginners, three unrated amateurs, two experts (ELO rating around 2,000 points) and a high-class international Grand Master.

Analysis of the results of this experiment showed a very significant ($p < 0.001$) correlation of reaction speed with chess ability: the Grand Master took around a third of the time to return a decision compared to the mean of the reaction times of the beginners.

Saariluoma noted that experienced players must perform at least some of the operations involved in the task more quickly than less experienced players, but did not predict which ones. It is known that a few of these processes are improved with practice, such as recognition of pieces (Saariluoma, 1984), speed of making moves in the mind's eye (Church & Church, 1977; Milojkovic, 1982); these have been addressed in our model (see below).

Whilst it is plausible that other cognitive processes involved may be improved through practice, we hypothesise that the greater relevant knowledge acquired by more experienced players should account for the main part of the remaining difference.

It is difficult to test this hypothesis on human subjects due to the obvious challenges of controlling for the amount of domain knowledge acquired and isolating the relevant processes. In order to investigate this hypothesis, the use of a cognitive model would be helpful in order to manipulate these factors directly.

A successful model should be able to demonstrate the superiority of experts over novices in the check detection task, and explain why.

## CHREST

CHREST (Chunk Hierarchy and REtrieval STructures) (Gobet et al., 2001) is a general-purpose cognitive architecture designed to simulate certain aspects of the human mind[1], including, to the extent that these have been measured or can

[1] For information on CHREST beyond what is presented here, the interested reader is referred to the CHREST website at http://chrest.info

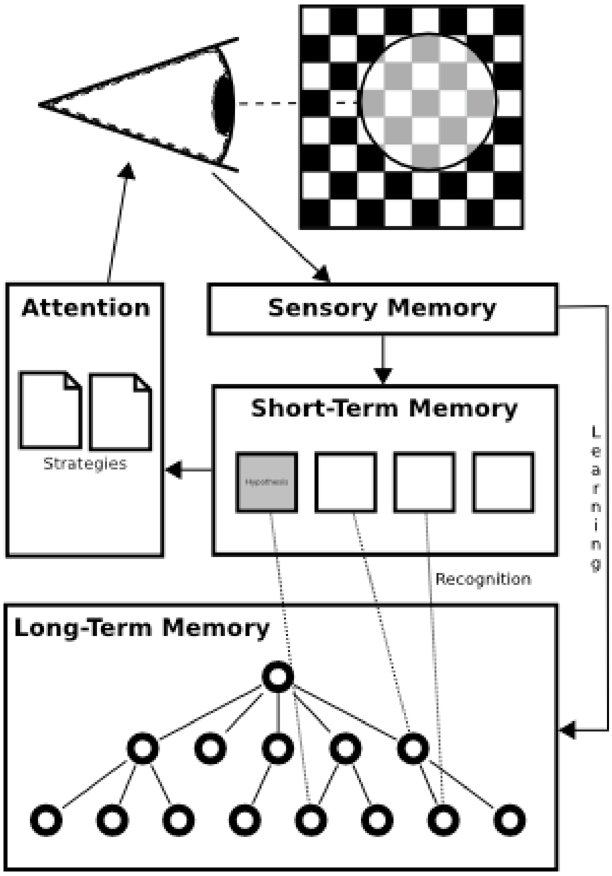


Figure 1: An overview of the main components of the CHREST cognitive architecture.

be estimated, its limitations (an important requirement of a model that aspires to simulate human cognition is that it not make use of any abilities in excess of that of a human (Simon, 1969)).

CHREST has previously been shown to be a successful model of the mind in domains as diverse as physics representation (Lane, Cheng, & Gobet, 2000), language acquisition (Jones, Gobet, & Pine, 2005; Freudenthal, Pine, Aguado-Orea, & Gobet, 2007), and ageing (Smith, Gobet, & Lane, 2007); however, the simulation of perception and memory in chess (de Groot & Gobet, 1996; Gobet & Simon, 2000; Gobet & Waters, 2003) has been CHREST's most studied application.

Figure 1 shows a top-level overview of the CHREST architecture. It simulates the main divisions of memory in humans as is generally agreed upon (Baddeley, 1990): a short-term memory (STM) store, and a long-term memory (LTM) store. In addition, it has an advanced perception/attention system.

CHREST's memory system is based on chunking theory (Chase & Simon, 1973), which holds that information in the human mind is stored as *chunks*. Chunks are discrete collections of *features* that have some meaning when grouped together. In the domain of chess, the features that make up chunks in CHREST are man-on-square combinations such as "White king on square g1". A chunk containing this feature, and representing a standard castled white king could be represented as the set: $\{$Kg1, Rf1, Pf2, Pg2, Ph2$\}$, where the first letter is the first letter of the piece's name.

CHREST's LTM is made up of a hierarchical network of these chunks. Its organisation is primarily tree-based, though the presence of semantic links adds a graph-like flavour. Knowledge is added to LTM through two main learning processes. When a new pattern is encountered, it is compared to previously-learnt chunks: if it does not match any known chunk, then a new chunk is created containing some of the new information (*discrimination*); if it does match a known chunk, then some of the information in the pattern is added to that chunk (*familiarisation*).

CHREST's STM has a capacity of up to four chunks. There is good evidence that this is the approximate STM capacity of young adult humans (Luck & Vogel, 1997; Cowan, 2001). These chunks are references to chunks held in LTM (again, as indicated by recent research (Gobet et al., 2001)).

Attention in CHREST is represented through simulated eye movements (this is a slight simplification of the human attention system, but this approach is relatively easy to simulate and its output can be verified against recorded human data).

CHREST's attention is directed through information previously learnt and added to LTM, and a set of heuristics. The basic heuristics, such as "look at the centre of the board", "look at objects grouped together", and "follow a potential move from an observed piece" guide the perception of basic patterns, which are incorporated into LTM as chunks.

As more information-rich chunks are acquired, this learnt information is used to guide the focus of attention. When an observed pattern is recognised as a previously-learnt chunk, a reference to the chunk is placed in STM, and this selected information may be used to provide a new focus. If a chunk referenced in STM is linked to another chunk in LTM, then CHREST's attention is directed towards locations containing objects in the linked chunk that are disjoint from the objects recorded in the recognised chunk. This process allows CHREST to focus on the distinguishing features of a scene.

Patterns, then, are perceived on the board according to previously-learnt chunks, and chunks are built up out of perceived patterns; this interplay between the learning cycle and the perception system results in complex emergent behaviour. In previous work (de Groot & Gobet, 1996), the eye movements generated by CHREST during a simulated presentation of a chess position have been shown to be comparable to those of Masters.

See (Lane, Gobet, & Smith, 2008) for more details of the attention system.

## CHREST Configuration

The version of CHREST used for these experiments was the 3.0 beta version. The code base of this version has been mostly rewritten from the 2.x version. It represents a substantial evolution of the model and a major step forward towards a full release of a complete CHREST 3. As well as being more flexible and better able to make use of modern computing technology, this version of CHREST has a number of new features: notably, it understands chess at a deeper level, includes a customised experiment framework which automatically performs and reports on sets of experiments, and has the ability to perceive and learn from full games rather than selected positions.

In order to simulate the variety of individuals employed in the human experiments, a series of CHREST subjects[2] with varying LTM sizes was produced: 20 each with a network size of one of 100, 1,000, 10,000, 100,000 nodes. (These different network sizes represented players of different skill level).

Most of the training of these subjects was carried out using a set of 10,000 games played during 2008 between players with ELO ratings of above 2000. Each subject was allowed to learn from the state of the game board at random intervals during simulated play-throughs of the games until the required network size was reached.

In addition, each subject was specifically trained on boards with a king and one other piece (all possible configurations of this type were produced to make up a training set). A total of 10% of each subject's LTM network was generated in this way, reflecting the fact that checking positions are very common in rapid and speed chess games, which most chess players use as a form of practice (Gobet & Campitelli, 2007). This figure is necessarily only an estimate of real-life behaviour due to a lack of empirical data at this time, but it was estimated in advance and not fitted to the model's result.

As in the Saariluoma experiment, 60 chess positions were generated for testing. A white king was placed on the board, along with a black queen, rook, bishop, or knight. For each position, the locations of the pieces were randomised, with the constraint that the king was placed on a square in which it was in check in half of the positions.

Timings played an important part in the experiments; time was one limiting factor for CHREST's perceptual cycle, and the time taken for CHREST to decide if a position contained a check was the main dependent variable in the experiments.

CHREST uses an internal clock which accumulates the processing times of simulated operations. These times are (where possible) taken from human experiments, or otherwise (where experimentation has not yet been possible), taken from sensible estimates (see (de Groot & Gobet, 1996) for details).

Unless otherwise noted, timings used were the standard timings which have evolved in CHREST:

- A constant 200 ms was added to all trials to simulate initial reaction to the stimulus, motor preparation, and motor response (i.e. pressing the button). Visual reaction times have been recorded as in the region 180 to 200 ms for university-age students (Brebner & Welford, 1980), though increasing with age (Welford, 1977).

- Saariluoma found, in a previous experiment (Saariluoma, 1984), that novices were slower than experts in recognising chessmen. The mean difference between the two groups was 57.1 ms; this value was added to the clock as time taken to recognise each piece for the 100 and 1,000-node network (this division was slightly arbitrary as it might be expected that the delay would be a gradient rather than binary, but we have no better data).

- From their analysis of experimental results, de Groot and Gobet (1996) proposed definite parameters for the time required to move pieces in the mind's eye. These parameters consisted of a *base* time, the time taken to begin making a move, and a *square* time, the time taken per square to move a piece. The first was estimated as 100 ms, and the second as 50 ms for experts, and 100 ms for novices. We have used these same values.

## Modelling Check Perception

We have described the domain of interest, that of the human process of perceiving and determining checks, and the general-purpose cognitive architecture that we are using to investigate it. Now we consider how to specifically adapt the model to the domain.

It has already been shown that the memorisation of chess positions under human constraints — see (de Groot, 1978) — can be improved through prior knowledge of chess positions. We propose that the process of determining whether a king is in check from another piece, given that the location and types of both pieces have been established, benefits from the presence in memory of previously encoded chunks of chess positions.

Our hypothesis for the superiority of experts over novices in detecting checks lies in chunking theory (Chase & Simon, 1973). Following previous work (Gobet & Jansen, 1994), we hypothesise that links are formed between the learnt visuo-spatial chunks and more abstract knowledge; for example, moves associated with the chunk, the goodness of the chunk in positional terms, and, of interest with respect to our particular domain, whether the chunk contains a check or not.

Our model, set up as above, simulates the experiment as follows. The simulated subjects were presented with a test position and allowed to perceive it until they had observed two pieces. Once this was achieved, they attempted to decide whether the king was being attacked by the other piece.

If the two pieces were recognised as a chunk already stored in their LTM, then the subject was assumed to be able to quickly (we have assumed 10 ms — the standard time taken to traverse an LTM link) identify whether the position was a check or not. Essentially, the subject would have exhibited automaticity (Shiffrin & Schneider, 1977).

[2] We use the term subject here to distinguish the computational instantiation of a model (complete with data) from the theoretical model

If no such chunk was recognised, then a simulated attempt to determine check was carried out, by 'moving' the non-king piece towards the king with simulated eye movements (previous work has demonstrated this proportionality to distance effect (Church & Church, 1977)), and checking this by 'moving' the king towards the other piece in a similar fashion. We assume that double verification of the check relation occurs here, but not when a chunk has been identified, as mentally moving pieces in the mind's eye is more likely to generate errors than when a pattern has been recognized (for a discussion of the difficulty of generating moves in the mind's eye and playing blindfold chess, see (Saariluoma, 1984), and (Campitelli & Gobet, 2005)).

## Experiment 1: Standard Perceptual Strategies

For the first experiment, our initial model made use of standard strategies to guide perception when studying each position: i.e. the use of LTM guidance, and fall-back general-purpose heuristics as described above and used in the training of the networks. These results are shown in table 1 and figure 2. The human data collected in (Saariluoma, 1984) are shown for comparison[3].

The results demonstrate some success of the approach in modelling the data ($r^2 = 0.92$): specifically, they show the required qualitative interaction of LTM size (acquired knowledge) with time, and are within around 200 ms of the times of the novice players. However, the results diverge from the human data considerably when considering performance of expert-level and above.

These strategies have previously been shown to be an accurate model of expert eye movements in perceiving scenes, but they clearly do not fully capture behaviour in this domain.

Table 1: Time taken to make a check perception decision as simulated by CHREST for players of different skill levels using standard perceptual strategies (Experiment 1).

| LTM Network Size (nodes) | Time Taken (ms) |
|---|---|
| 100 | 1,705 |
| 1,000 | 1,403 |
| 10,000 | 1,301 |
| 100,000 | 1,068 |

## Experiment 2: Simplified Perceptual Strategy

Following the results of the first experiment, we hypothesised that players are using their meta-knowledge about the problem to re-orient their perceptual strategies. As the model consistently overestimated the time taken, we suspected that the perceptual strategies used by CHREST were too involved and that humans used a simpler strategy.

Our revised model was that the subject would automatically perceive a man on the board using far peripheral vision and direct their attention towards it. The subject would then make use of their near peripheral vision (set at $\pm 2$ squares from the focal point) to recognise another piece if one was in range. If no piece was in range, the player would detect the other piece using their far peripheral vision, and refocus on that point following a saccade (thus, making one, or a maximum of two, eye fixations; in the previous experiment, the focus could be directed towards empty squares).

[3]The exact human data were not available and so have been read from the graph supplied in (Saariluoma, 1984)

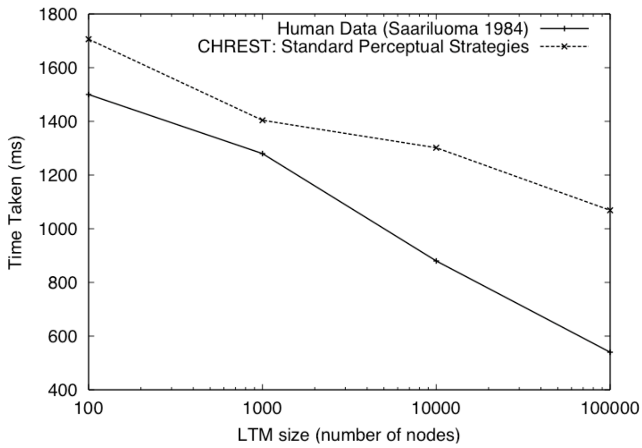


Figure 2: Time taken to make a check perception decision as simulated by CHREST for players of different skill levels using standard perceptual strategies (Experiment 1). The human data are shown for comparison.

The results of re-running the experiment using this strategy are shown in table 2 and figure 3. This time the results are a significantly better fit to the human data ($r^2 = 0.94$), again showing the qualitative interaction, but matching the data quantitatively to within 200 ms at worst. In this experiment, however, the results better match the data for advanced players rather than novices.

Table 2: Time taken to make a check perception decision as simulated by CHREST for players of different skill levels using standard perceptual strategies (Experiment 2).

| LTM Network Size (nodes) | Time Taken (ms) |
|---|---|
| 100 | 1,320 |
| 1,000 | 1,010 |
| 10,000 | 883 |
| 100,000 | 606 |

# Discussion of Results

Before discussing the results, we note that there are some limitations to the study and suggest some other reasons for caution in interpreting the results.

We have assumed above that our choices of four network sizes — {100, 1,000, 10,000, 100,000} — correspond to Saariluoma's categorisation of his subjects — {Fourth

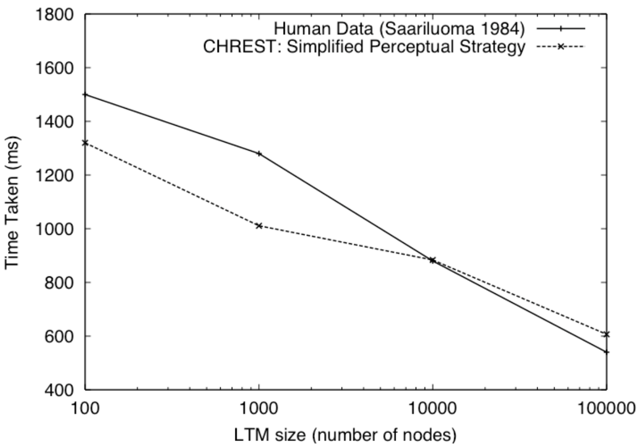


Figure 3: Time taken to make a check perception decision as simulated by CHREST for players of different skill levels using standard perceptual strategies (Experiment 2). The human data are shown for comparison.

Class, Second, Class, Experts, Grand Masters} — but this is not clear. For obvious reasons it is not possible to directly measure the number of chunks learnt by a human subject, so we used a logarithmic progression as an approximation. 100 chunks is probably too many for a beginner who is still learning how the pieces move (though it makes no difference to the result as no chunks were recognised by the 100-node networks) and estimates of the number of chunks learnt by a Grand Master differ, from 100,000 (Simon & Gilmartin, 1973) to 300,000 (Gobet & Simon, 2000) (but this may not be an issue as we argue below that larger networks will probably not show much relative improvement).

The estimate that 10% of a subject's training is on endgame positions is difficult to verify. However, as noted earlier, it is known that players play a large number of speed chess games, where check situations are frequent, and the proportion was estimated in advance of the experiment, so we believe that the figure is reasonable pending other evidence.

A small number of errors were produced by the human subjects: a mean of 3.0%, with a maximum of 4.1% by the experts. CHREST is theoretically able to produce errors (for example, by over-generalising learnt information), but none were produced in these simulations. This may be considered a weakness of the model, but given the proportion of errors made by humans, and that a number of these may have been due to errors of attention (e.g. pressing the wrong key due to fatigue), we do not think this is a serious drawback.

Despite these considerations, we find the results good evidence for our hypothesis. We have proposed a model of how humans carry out simple check detection and found that, with a revision and accepting certain assumptions, it explains the human data well, both quantitatively and qualitatively.

Our revised model shows poorer performance in modelling the perception of weaker players. This may be natural variability, given Saariluoma's small sample size, but we also consider other possible reasons:

Our first model may have been partly right, and though stronger players do use the more efficient, simplified, perceptual strategy described in our revised model, weaker players use (a subset of) the unnecessarily complicated strategies used for perceiving a regular game position.

Alternatively (or in addition), there is some evidence (Reingold, Charness, Pomplun, & Stampe, 2001) that stronger players make better use of their peripheral vision to detect pieces, suggesting that we may have allowed weaker players too much ability in our revised model.

Also, weaker players may spend more time checking their decisions. We have assumed that a "double check" is carried out (checking the relationship between the position of both perceived pieces), but weaker players may find it necessary to make additional checks. It would be expected that stronger players would not feel the need to do this due to their improved confidence in their own ability.

Finally, there may be additional mental processes involved with weaker players which we have not considered. For example, absolute beginners may spend some time trying to remember how each piece moves.

Looking forward, our theory makes predictions that can be tested. Most obviously, players' eye movements could be recorded whilst carrying out this task to determine if our theory of how attention is directed (i.e. very simply and directly) is correct.

Our theory, that chunks are linked to further knowledge, including information about whether a chunk includes a check or not, also leads to some predictions.

First, there should be increased intra-subject variability across different positions compared to the "general exercise" hypothesis of several different mental processes being improved as expertise is acquired: general-purpose processes should not be affected by the specifics of an individual position.

Second, there should be a ceiling to performance on the task. The largest network we tested was 100,000 nodes, but there are only 16,128 separate possible positions containing only a king and one of {queen, rook, bishop, knight} of the opposite colour. Our imposed end game-specific practice of 10% (estimated, but seeming to produce a good match of the human data) of a 100,000 node network covers the majority of these positions. If our theory is correct, performance on this task should rapidly tail off above Grand Master level since there will be fewer additional novel chunks to acquire.

## Conclusions and Future Work

We have proposed a hypothesis of how humans perceive and make decisions on checks in simple check positions and from this produced a model that successfully reproduces the experimental human data.

This theory may have wider implications in terms of chunking theory. We have suggested that chunks are linked to extended information — specifically, information about whether the king is in check or not. This theory (if correct) raises a number of questions about the extent and types of information that may be linked to visuo-spatial chunks. Chunks could, for example, be linked to additional semantic information about their strategic value, their relationship to other chunks, or a verbal description. Based on earlier work (Gobet & Jansen, 1994), we are currently attempting to expand this theory by investigating how move sequences in chess may be learnt and attached to visual chunks in a similar manner.

Another, more direct, way to build on this work is to consider checks involving multiple pieces, for example in mid game chess positions. More complicated perceptual strategies would undoubtedly be involved.

Finally, in order to successfully model the human data, we have had to modify the perceptual strategies used, following the assumption that this behaviour would be controlled by conscious processes. Whilst this is a reasonable assumption backed by evidence, it required human intervention; ideally, the model would be able to alter its own behaviour in this way, controlling what information entered STM and directing its own perceptual strategies.

## Acknowledgment

This research was funded by the Economic and Social Research Council under grant number RES-000-23-1601.

## References


Baddeley, A. (1990). *Human memory. Theory and practice.* Boston: Allyn & Bacon.

Brebner, J., & Welford, A. (Eds.). (1980). *Reaction times*. New York: Academic Press.

Campitelli, G., & Gobet, F. (2005). The mind's eye in blindfold chess. *European Journal of Cognitive Psychology*, *17*, 23–45.

Chase, W., & Simon, H. (1973). The mind's eye in chess. In W. Chase (Ed.), *Visual information processing* (pp. 215–281). New York: Academic Press.

Church, R., & Church, K. (1977). Plans, goals, and search strategies for the selection of a move in chess. In P. Frey (Ed.), *Chess skill in man and machine* (pp. 131–156). New York: Springer Verlag.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–114.

de Groot, A. D. (1978). *Thought and choice in chess* (2nd ed.). The Hague: Mouton Publishers.

de Groot, A. D., & Gobet, F. (1996). *Perception and memory in chess: Heuristics of the professional eye*. Assen: Van Gorcum.

Freudenthal, D., Pine, J. M., Aguado-Orea, J., & Gobet, F. (2007). Modelling the developmental patterning of finiteness marking in English, Dutch, German and Spanish using MOSAIC. *Cognitive Science*, *31*, 311–341.

Gobet, F., & Campitelli, G. (2007). The role of domain-specific practice, handedness and starting age in chess. *Developmental Psychology*, *43*, 159–172.

Gobet, F., & Jansen, P. (1994). Towards a chess program based on a model of human memory. In H. van den Herik, I. Herschberg, & J. W.Uiterwijk (Eds.), *Advances in computer chess* (Vol. 7). Maastricht: University of Limburg Press.

Gobet, F., Lane, P. C. R., Croker, S. J., 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. (2000). Five seconds or sixty? presentation time in expert memory. *Cognitive Science*, *24*, 651–682.

Gobet, F., & Waters, A. (2003). The role of constraints in expert memory. *Journal of Experimental Psychology: Learning, Memory Cognition*, *29*, 1082–1094.

Jones, G. A., Gobet, F., & Pine, J. M. (2005). Modelling vocabulary acquisition: An explanation of the link between the phonological loop and long-term memory. *Journal of Artificial Intelligence and Simulation of Behaviour*, *1*, 509–522.

Lane, P. C. R., Cheng, P. C.-H., & Gobet, F. (2000). CHREST+: Investigating how humans learn to solve problems with diagrams. *AISB Quarterly*, *103*, 24–30.

Lane, P. C. R., Gobet, F., & Smith, R. Ll. (2008). Attention mechanisms in the CHREST cognitive architecture. In L. Paletta & J. Tsotsos (Eds.), *Proceedings of the 5th international workshop on attention in cognitive science* (pp. 207–220). Graz: Joanneum Research.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281.

Milojkovic, J. (1982). Chess imagery in novice and master. *Journal of Mental Imagery*, *6*, 125–144.

Reingold, E., Charness, N., Pomplun, M., & Stampe, D. (2001). Visual span in expert chess players: Evidence from eye movements. *Psychological Science*, *12*, 48–55.

Saariluoma, P. (1984). *Coding problem spaces in chess: A psychological study* (Vol. 23). Helsinki: The Finnish Society of Sciences and Letters.

Shiffrin, R., & Schneider, W. (1977). Controlled and automatic human information processing: 2. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*, 127–190.

Simon, H. (1969). *The sciences of the artificial*. Cambridge, MA: MIT Press.

Simon, H., & Gilmartin, K. (1973). A simulation of memory for chess positions. *Cognitive Pyschology*, *5*, 29–46.

Smith, R. Ll., Gobet, F., & Lane, P. C. (2007). An investigation into the effect of ageing on expert memory with CHREST. In *Proceedings of the seventh UK workshop on computational intelligence.*

Welford, A. (1977). Motor performance. In *Handbook of the psychology of aging* (pp. 450–496). New York: Van Nostrand Reinhold.