Some Shortcomings of Long-Term Working Memory

Fernand Gobet

ESRC Centre for Research in Development, Instruction and Training
Department of Psychology
University of Nottingham
Nottingham NG7 2RD
England

(0115) 951 5402

frg@psyc.nott.ac.uk

Author's note

Abstract

Within the framework of their long-term working memory theory, Ericsson and Kintsch (1995) propose that experts rapidly store information in long-term memory through two mechanisms: elaboration of long-term memory patterns and schemas and use of retrieval structures. They use chess players' memory as one of their most compelling sources of empirical evidence. In this paper, I show that evidence from chess memory, far from supporting their theory, limits its generality. Evidence from other domains reviewed by Ericsson and Kintsch, such as medical expertise, is not as strong as claimed, and sometimes contradicts the theory outright. I argue that Ericsson and Kintsch's concept of retrieval structure conflates three different types of memory structures that possess quite different properties. One of these types of structures—generic, general-purpose retrieval structures—has a narrower use than proposed by Ericsson and Kintsch: it applies only in domains where there is a conscious, deliberate intent by individuals to improve their memory. Other mechanisms, including specific retrieval structures, exist that permit a rapid encoding into long-term memory under other circumstances.

Some Shortcomings of Long-Term Working Memory

Ever since the seminal work of De Groot (1946/1978) on chess expertise, psychologists have attempted to understand the mental mechanisms and structures that allow some people to excel in domains that seem to overtax humans' limited cognitive system. One of the most successful attempts has been Chase and Simon's (1973) chunking theory, which accounts both for data from classical laboratory experiments and for empirical evidence on expert behaviour. Recently, and partly as a result of the perceived limitations of the chunking theory, Ericsson and Kintsch (1995) have proposed a new comprehensive theory—the long-term working memory (LT-WM) theory. It is the goal of this paper to discuss and evaluate the explanatory power of LT-WM in general, and of the construct of <u>retrieval structure</u>, one of its key ingredients, in particular.

The Chunking Theory and the Skilled Memory Theory

De Groot (1946/1978) was mostly interested in the way chessplayers organise their thoughts when trying to decide which move to play next. Somewhat to his surprise, his analysis of verbal protocols did not detect differences in the macro-structure of thought (e.g., depth of search, number of nodes searched, etc.), except that better players obviously found better moves than weaker players. The rapidity with which his best players were able to home in on important aspects of a position led De Groot to the hypothesis that the key to expertise lies in a vast knowledge base, made accessible by highly tuned perceptual mechanisms. To test his hypothesis, he presented chess positions for a short amount of time (from two to fifteen seconds), and, as predicted,

found a huge difference in recall performance between masters and weaker players.

De Groot's theoretical explanation was rather vague, however, and one had to wait until 1973 to have a precise, information-processing model, with Chase and Simon's chunking theory, based on the EPAM theory (Elementary Perceiver and Memoriser; Feigenbaum & Simon, 1962, 1984). Chase and Simon (1973), proposed that chessplayers, as well as other experts, acquire a large number of chunks (familiar units denoting perceptual patterns) through practice and study. Chunks are accessed by sorting stimuli through a discrimination net, where various perceptual features are tested (see Feigenbaum & Simon, 1984, for details), and act as the conditions of productions (Newell & Simon, 1972). In the case of chess, chunks evoke information such as plausible moves, potential plans, or evaluations of portions of the chess position at hand. In addition, according to Chase and Simon, general cognitive parameters, such as the time to learn a chunk or the number of chunks that can be held in short-term memory (STM), do not differ between experts and non-experts, and place stringent limits on what operations can be carried out. For example, the chunking theory proposes that it takes about eight seconds to learn a new chunk, and that only about seven chunks can be held in STM. Chess masters can thus recall a briefly presented position almost perfectly because they can recognise more and larger chunks than weaker players, and they can find better moves because chunks give rapid access to key information that may be elaborated by further look-ahead. Mutatis mutandis, Chase and Simon's theory can be, and has been, applied to other domains of expertise.

This simple model has generated a wealth of empirical studies, which have confirmed most of its predictions, but have also uncovered two main weaknesses. First, information seems to be transferred into long-term memory (LTM) faster than proposed by the chunking theory. For example, Charness (1976) has found that interpolating a task between the presentation of a position and its recall affects performance only slightly. Second, it is hard to see how such a small memory capacity allows experts to carry out complex tasks, such as search through the problem space. This point is made very strongly in Ericsson and Kintsch (1995) and is buttressed by the fact that current production systems such as ACT (Adaptive Control of Thought; Anderson, 1983; Anderson & Lebière, 1998) and Soar (State, Operator, And Result; Newell, 1990) require a working memory much larger than a half dozen chunks.

An important alternative to the chunking theory was proposed by Chase, Ericsson and Staszewski (Chase & Ericsson, 1982; Ericsson & Staszewski, 1989; Staszewski, 1990), who supported their theory by detailed studies of mnemonists and experts in mental calculation. Their skilled memory theory explains experts' remarkable memory and problem-solving abilities through three principles: (a) Information is encoded with numerous and elaborated cues related to prior knowledge; (b) Experts develop a retrieval structure, that is a LTM structure for indexing material in LTM. As a classical example of a retrieval structure, take the method of loci, in which one uses a general, predetermined set of locations. During presentation of the material to memorise, associations are made between the locations and the items to learn. At recall, the locations are used as retrieval cues to access the items to be

recalled; and (c) Time required by encoding and retrieval operations decreases with practice.

Chase and Ericsson (1982) note that it takes extensive practice to develop and use a retrieval structure. Typically, under rapid presentation, the structure can be used to successfully encode only one type of material, say digits, without transfer to other material. Finally, the retrieval structure must be activated before the material is presented.

The Long-Term Working Memory Approach

Recently, the skilled memory theory has been extended into the long-term working memory (LT-WM) theory (Ericsson & Kintsch, 1995). Using data from tasks such as memory for menu orders, the digit-span task, mental multiplication, mental abacus calculation, chess, medical expertise, and text comprehension, Ericsson and Kintsch (1995) present compelling empirical evidence that experts in various domains are able to encode information into LTM faster than was proposed by traditional models of human memory, including the chunking theory. The core of their long-term working memory theory is that "cognitive processes are viewed as a sequence of stable states representing end products of processing" and that "acquired memory skills allow these end products to be stored in long-term memory and kept directly accessible by means of retrieval cues in short-term memory [...]" (Ericsson & Kintsch, 1995, p. 211). Encoding occurs through a retrieval structure and/or through knowledge-based associations connecting items to other items or to LTM patterns and schemas, which allows for an integrated representation of the information in LTM. The task demands on memory dictate which encoding method will be used.

A key element in the LT-WM theory, which distinguishes it from most other theories of memory, is that experts develop, through practice and study, retrieval structures for the task domain. Retrieval structures are "a set of retrieval cues [that] are organized in a stable structure" (Ericsson & Kintsch, 1995, p. 216). The best empirical evidence for these structures is offered by the detailed study of SF and DD, two experts in the digit-span task (Chase & Ericsson, 1982; Staszewski, 1990). It is proposed that SF and DD store groups of digits and additional semantic information in a hierarchical retrieval structure. Analysis of reaction times and of verbal protocols, as well as experimental manipulations, provides strong support for the presence of retrieval structures within these tasks. As stated by the theory, in addition to retrieval structures, LTM elaborations of previously stored semantic memory facilitate the encoding of information by adding redundant cues, and the relative roles played by the retrieval structure and LTM learning vary from task to task.

The main thrust of Ericsson and Kintsch's paper is to propose that the basic tenets of the skilled memory theory (Chase & Ericsson, 1982), including the presence of retrieval structures, are generalizable across a wide variety of tasks. Thus, the LT-WM theory goes beyond the more limited claim that the concept of retrieval structures is useful only in tasks where subjects show a deliberate and conscious attempt to improve their memory through strategies ("mnemonics") and where the scheme for subjective organisation of the material to be recalled is known to subjects beforehand. As a consequence, the LT-WM theory conflicts with traditional accounts of working memory, which stress its transitory storage capacity. If correct, the LT-WM theory would be a major step in our understanding of human cognition. It puts

together, in a unified framework, empirical results from domains that are normally treated separately: classical experimental studies on memory, as well as research on expertise, text comprehension, and problem solving.

In this paper, I will show that this claim of generality is exaggerated. Although there is no doubt that most experts have rapid (although less impressive than SF or DD's) access to LTM, the concept of retrieval structure, which plays a key role in the LT-WM theory, suffers from a number of weaknesses: the concept is not sufficiently well specified to allow precise predictions to be made (this also applies to the LT-WM theory in general); it is used inconsistently; and it has received weaker empirical support than claimed by Ericsson and Kintsch.

I will start by considering the lack of specificity of LT-WM and discussing some inconsistencies in the way Ericsson and Kintsch (1995) use the concept of retrieval structure. I will then analyse in detail the application of LT-WM to the domain of chess expertise, a model task in the study of expertise (Charness, 1992), which has the advantages of a long scientific tradition going back to Binet (1893/1966), and of offering a large amount of empirical data. Since Ericsson and Kintsch (1995) note that "research on planning and memory of chess positions offers some of the most compelling evidence for LT-WM" (p. 238), the choice of this domain is adequate to test their theoretical claims of generality. Then, I will briefly discuss the application of LT-WM to some of the other domains of expertise reviewed by Ericsson and Kintsch. Finally, I will compare the LT-WM account with two recent computational theories of expertise, one devoted to the digit-span task (EPAM-IV; Richman, Staszewski & Simon, 1995) and the other to chess expertise (template theory; Gobet & Simon, 1996b; 1998; in press), in order to

highlight the similarities and the differences in the way these theories use the concept of retrieval structure.

Lack of Specificity of the Theory and Inconsistency in Using the Concept of Retrieval Structure

One serious difficulty with understanding, applying, and evaluating LT-WM is that Ericsson and Kintsch (1995) provide no definition of the concepts of patterns and schemas, which are key elements of their theory and which both have a large variety of meanings in cognitive psychology. As a consequence, the reader has to use his or her own definition of patterns and schemas, which may or may not correspond to Ericsson and Kintsch's. Another difficulty is that the theoretical properties that they assign to the concept of retrieval structure are not consistent across domains. For example, there are important differences in Ericsson and Kintsch's use of retrieval structure in their treatment of text comprehension as compared with their use of that concept in other domains. Recall that Ericsson and Kintsch (1995, p. 220) propose that LT-WM includes two types of retrieval mechanisms: "cue-based retrieval without additional encodings," which refers to retrieval structures, and "cue-based retrieval with an elaborated structure associating items from a given trial or context," which refers to knowledge-based associations. Recall also that Ericsson and Kintsch (p. 216) define a retrieval structure as a set of retrieval cues organised in a stable structure. Their nomenclature changes when dealing with text comprehension, where they refer to two different sources of retrieval structures: domain knowledge and episodic text structure (pp. 230-231). Inconsistencies are apparent when the previous definition of retrieval structures is compared with the episodic text structure, which is generated case

by case, based on several levels of discourse (surface features, propositions, macropropositions, semantic relations, rhetorical relations; Ericsson & Kintsch, 1995, p. 230). This structure does not meet the criterion of stability mentioned earlier, as it becomes harder to use as subsequent sentences are processed. In addition, this structure is rapidly created (in a matter of seconds) as a sentence is processed, while it takes several months to develop the retrieval structures described in domains of expertise such as the digit-span task. I will come back to the difficulties associated with the lack of specificity of LT-WM in the next section, after describing the LT-WM account of chess expertise.

LT-WM Account of Chess Expertise

Building on previous research mentioned in Ericsson and Staszewski (1989), Ericsson and Kintsch (1995) propose that skilled chess players use a hierarchical retrieval structure corresponding to the 64 squares of the chess board. Thanks to this structure, which both relates pieces to each other and associates individual pieces to their respective locations, a position is represented as an integrated hierarchical structure, which allows a rapid encoding into LTM. LTM itself stores, among other things, patterns and schemas, into which new information can be encoded rapidly, independently of the use of the retrieval structure. It is proposed that the retrieval structure offers an explanation of how chess masters can plan and evaluate alternative sequences of moves.

With the explicit goal of supporting the hypothesis that chess masters use a retrieval structure, Ericsson and Kintsch (1995, pp. 237-238) review several pieces of evidence: results from recall experiments where the random

chess positions lack meaningful configurations; results showing that skilled chess players can rapidly access the location of pieces within a memorised chess position; and results showing that masters can mentally manipulate and update the contents of a position when playing blindfold chess. In addition, Ericsson and Kintsch (1995, p. 214-215 and 237-238) discuss a few other chess experiments. Of these, I will focus on: interference due to concurrent memory tasks, role of piece grouping during presentation, and recall of multiple boards.

Vagueness of LT-WM Account for Chess Expertise

Although chess is presented as providing some of the strongest support for LT-WM, the theory is stated in rather general terms and leaves many crucial parameters and necessary mechanisms unspecified: What is the hierarchical organisation of the retrieval structure? What type of information (pieces, chunks, or higher-level schemas) can be encoded at the various levels of the hierarchy? How long does it take to encode a retrieval cue? Are cues subject to some decay?[1] Are cues for all the "end products of processing" encoded in the retrieval structure, or is the encoding probabilistic? Is encoding under strategic control? And, for that matter, what counts as "end products of processing" within the chess domain? While it is true that Ericsson and Kintsch's goal was not to offer a detailed theory of chess expertise but to show how LT-WM could account for data from a large variety of domains, it is also true that without the kind of specifications just mentioned, it is simply impossible to evaluate the LT-WM explanation of chess expertise—or for other domains.

In addition, most of the experiments reported in their paper use rather long presentation times (typically, 2 s per move or 2 s per piece), and Ericsson and Kintsch are not explicit about what mechanisms are used with experiments showing recall with short presentation times, say, less than 10 s for the entire position (typically 25 pieces), such as the standard De Groot (1946/1978) recall task. It is clear from a previous publication (Ericsson & Staszewski, 1989) that the retrieval structure is also assumed to be used efficiently with short presentation times, but the reader is left in the dark as to what really happens in this case. This is a weak point of the theory, as most evidence in chess research has been gathered with such short presentation times.

It has been known since De Groot (1946/1978) that strong players can recall a game position almost perfectly with a presentation time of 5 s and still get high percentages of recall with presentation times as low as 2 s (for additional data on very short presentation times, see Gobet & Simon, in press). How does LT-WM account for these results? A first interpretation is that pieces are encoded into the squares of the retrieval structure. This interpretation rapidly runs into several problems, however. Two examples will suffice. First, the encoding times per unit of information (piece) must be very short, much shorter than in the other tasks reviewed by Ericsson and Kintsch (1995). With all other domains of expertise discussed by Ericsson and Kintsch, the presentation time per unit of information (e.g., one digit or one menu item) is at least 1 s, and experts therefore have at least 1 s to encode each unit. In chess, where masters can recall almost perfectly a position containing 25 pieces with a presentation time of 5 s, we have to assume that individual units of information (the chess pieces) are encoded into the

retrieval structure very fast, in the order of about 200 ms (5 s divided by 25).[2]
Second, recall of random positions is below what would be predicted by this
interpretation of the theory: even if masters maintain a modest superiority over
weaker players with presentation times up to 10 s (Gobet & Simon, 1996a),
their performance is poor (on average, 5.5 pieces out of 25), and is not above
what would be predicted if masters were only storing information in STM, as
shown in the computer simulations discussed in Gobet (1998b).

A more plausible interpretation is that, rather than encoding pieces into
squares, experts encode schemas and patterns into higher levels of the
hierarchical retrieval structure. Given such a mechanism, the LT-WM theory
is consistent with masters' performance with rapid presentation times (for
example, random positions are harder than game positions because it is less
likely that patterns or schemas are retrieved). With this interpretation,
however, all the explanatory power of the theory rests on patterns and
schemas, and the concept of retrieval structure is not necessary, a traditional
limited-size STM being sufficient (cf. Gobet & Simon, 1996b). Finally,
encoding cues for patterns and schemas into the retrieval structure raises one
important question, which Ericsson and Kintsch do not address: By what
mechanism is information contained in these LTM entities "unpacked" at
lower levels of the retrieval structure, allowing masters to rapidly access the
location of each piece?

Empirical Data from Chess Research Taken as Support for the LT-WM Theory

I now review the empirical evidence from chess memory research given by
Ericsson and Kintsch (1995) as supporting their position. As I will show, this

evidence, far from supporting LT-WM, raises serious difficulties for the concept of retrieval structure, and, as a consequence, for the LT-WM theory in general.

Random Positions with Long Presentation Time

Random material is crucial in evaluating the hypothesis of retrieval structures, for, with experiments using meaningful material, it is difficult to tease apart the role of such structures and that of patterns and schemas. According to Ericsson and Kintsch (1995, p. 237), "skilled chess players are able to encode and store the locations of individual chess pieces of a chess position in the absence of meaningful configurations of chess pieces." This comment is surprising, because empirical evidence about random positions is devastating for the LT-WM theory: with short (see above) and long presentation times, players of all skill levels have huge difficulties with this type of position.

With visual presentation of the whole board for 60 s, masters do not recall more than 70% of the pieces correctly (Gobet & Simon, in press). With an auditory, piece-by-piece average presentation of about 50 s, at a rate of one piece every 2 s, masters do not recall more than around 60% (Saariluoma 1989). In addition, Saariluoma (1989), again using dictation of pieces, has shown that masters recall little from four random positions presented in sequence (around 10% per position), while they perform relatively well with game positions (around 60% correct per position). Interestingly, masters' recall performance with random positions is roughly what is predicted by one of the traditional theories dismissed by Ericsson and Kintsch (1995)—the Chase and Simon (1973) chunking theory, which imposes rather strong limits on memory capacity and on learning rate.[3]

I will focus in my analysis on the auditory presentation, as the results with visual presentation were not available to Ericsson and Kintsch. In their instantiation of the LT-WM theory for experiments where positions are dictated, Ericsson and Kintsch state that

"If, on the other hand, chess experts had a retrieval structure corresponding to a mental chess board, they could store each piece at a time at the appropriate location within the retrieval structure. After the end of the presentation the experts would be able to perfectly recall the entire position if the presentation rate had been slow enough." (1995, p. 237, italics added).

In addition, Ericsson and Kintsch mention that, with game but not with random positions, meaningful patterns of relations between pieces can be encoded, allowing the position to be stored as more integrated structures. This ability to find higher-order relations in game positions explains why such positions are easier to recall than random positions.

This explanation, while accounting for the superiority in the recall of game positions over random positions, is unsatisfactory, however, when one keeps in mind that Ericsson and Kintsch reviewed Saariluoma's (1989) experiments to show that "the ability to store random chess positions provides particularly strong evidence for the ability to encode individual chess pieces into the retrieval structure" (Ericsson and Kintsch, 1995, p. 237; italics added) and that recall for random positions is far from perfect even with masters (not more than about 60%, that is about 15 pieces, after a 50-second presentation). To begin with, the retrieval structure allows the encoding of individual pieces and, therefore, should be useful even for recalling random positions. Furthermore, there are at least as many relations of attack, defence, colour,

and proximity in random positions as in game positions (Chase & Simon, 1973), which should allow experts to encode redundant cues, such as relations between pieces,[4] into the hierarchical retrieval structure, in particular with presentation times of 50 s for the entire position. In addition, the other mechanism provided by LT-WM—the ability to rapidly generate new associations of LTM patterns and schemas—should be able to operate with random positions as well, the term "pattern" being sufficiently vague to include small chunks of pieces such as found in random positions. Note, finally, that the retrieval structure used with SF and DD in the digit-span task allows them to remember sequences of digits that are essentially random. That SF and DD are able to reliably and rapidly recode random items into meaningful chunks, while chess masters are not, seems to point to a key difference between the cognitive processes and the task environment of the digit-span and chess experts. Ericsson and Kintsch suggest that masters' relatively low performance with random chess positions is accounted for by the difficulty of encoding them into "new distinct structures" in LTM.[5] This explanation undermines their claim that recall of random positions strongly supports the hypothesis of a retrieval structure, as it rests on the absence of schemas, and not on positive properties of the retrieval structure.

The LT-WM theory actually seems to be caught in a dilemma. On the one hand, it proposes a powerful retrieval structure and the capacity to rapidly make new associations into LTM. On the other hand, data with random positions show that chess experts have a harder time remembering random positions than would be predicted by the theory. Ericsson and Kintsch's solution to this dilemma is to propose that information can be encoded as patterns and schemas into the retrieval structure with game positions, but not

with random positions. This solution does explain the superiority of recall with game positions over that with random positions, but does not explain why, in spite of the powerful mechanisms and structures associated with LT-WM, masters do not recall random positions better. Nor does this solution provide any evidence that the recall of random positions supports the hypothesis of a chess retrieval structure, although this was the explicit goal of reviewing data on the recall of dictated random positions.

Mental Manipulation and Updating of the Board

Citing unpublished experiments by Ericsson and Oliver (1984; see Ericsson and Staszewski, 1989, for a summary), Ericsson and Kintsch (1995) propose that the retrieval structure hypothesis accounts for the ability of a master to rapidly access the information of individual squares of the chessboard. With the qualification, already mentioned above, that they do not specify mechanisms describing how information can be transferred from the high levels of the retrieval structure to the square level, their argumentation appears sound. They also refer to experiments by Saariluoma (1991) using blindfold chess, and propose that the retrieval structure hypothesis offers an explanation for chess masters' ability to play without external perceptual support at a level close to their normal skill. This part of their argument requires several comments.

To begin with, one should mention that no formal study has established to what extent masters can play blindfold chess at a level close to that in normal chess. Holding (1985) actually suggested that the quality of play in the two situations differs radically, blindfold games being characterised by many more blunders than normal games. However, anecdotal evidence also shows that some blindfold games can be of rather good quality.

We therefore need to look at laboratory studies to test Ericsson and Kintsch's (1995) claims about masters' ability to represent a position during a blindfold game by using a retrieval structure.

Saariluoma (1991) has carried out an experiment which directly bears on this question and clearly illustrates the difficulty of playing blindfold chess. He dictated a game at a rate of one half-move (move for White or Black) every 2 s, and asked players to indicate the location of the pieces after 30 half-moves. He found that masters could recall more than 90% of the position when the moves were taken from an actual game or when the moves were random, but legal. However, when the game consisted of (possibly) illegal moves, the recall dropped to only around 20%, which was barely better than weak players' recall. Using the chunking theory (Chase & Simon, 1973) as a framework, Saariluoma's (1991) explanation was that random legal games drift only slowly into chaotic positions where few chunks can be recognised, while random illegal games move more rapidly into such positions. It is unclear what explanation the LT-WM theory offers for the differential recall of random legal and random illegal games, because in both cases, playing a move produces changes on the board that can be equally well encoded into the retrieval structure or into LTM elaborations. It is also unclear how the LT-WM theory could incorporate Saariluoma's explanation based on chunking without drastically reducing the role played by the retrieval structure in chess players' memory. As was the case with the recall of random positions with long presentation time, Saariluoma's (1991) experiment fails to provide positive evidence for the retrieval structure that is not confounded by the presence of patterns and schemas.

Interference Due to a Concurrent Memory Task

Saariluoma (1992) has shown that, in a move generation task, a (concurrent) articulatory suppression task did not impair performance, but that a visuo-spatial task did. Saariluoma (1991) found the same pattern of results in a task where subjects had to mentally construct the correct positions from the moves dictated by an experimenter (see Robbins et al., 1996, for similar results). Finally, he found that interfering tasks (both visuo-spatial and articulatory) had no effect when they were carried out after a sequence of moves had been dictated. While these results are compatible with the LT-WM theory, it is unclear why they should "clearly implicate LT-WM in the maintained access to the updated chess positions" (Ericsson & Kintsch, p. 238). The dictation time per move was long enough in these experiments (at least 2 s per move) to allow relatively slow encoding into LTM (many sequences of moves in master games are standard, and thus may be coded as chunks), and traditional models of chess memory, most of which stress the visual code used by chess players (e.g. Chase & Simon 1973), would predict the same result. Finally, the result that a concurrent task impairs learning more than a posterior task is consistent with all current theories of memory.

Piece Grouping during Presentation

In their summary of expert memory in domains other than text comprehension, Ericsson and Kintsch (p. 238) note that

> "The strongest evidence for retrieval structures concerns the ability of experts to independently store pieces of information when they are presented out of their normal context in scrambled order. After such a presentation, experts in medicine and chess and a waiter (JC) were able to recall all of the information in an order reflecting its typical meaningful organization in the retrieval structure."

However, experimental results—some of them reviewed in their paper—tell another story. Charness (1976), as well as Frey and Adesman (1976), has shown that pieces presented at a rate of about 2 s per piece are better retained when they are presented by chunks similar to the ones identified by Chase and Simon (1973a) than when they are presented by columns or randomly; this result is valid for both verbal and visual presentation. Saariluoma (1989) found a similar result between an ordered reading (White pieces are dictated first, starting from Pawns, and then from King to Knight in decreasing chess value) and a random-order reading. According to the LT-WM framework, the presence of a powerful retrieval structure should allow experts to independently store units of information regardless of the type of grouping (by chunks, by columns, or in scrambled order) during presentation. The results reported above show that chess experts do not have this ability and, therefore, directly contradict one important prediction of the LT-WM theory. (As discussed below, this prediction is also incorrect with medical expertise.)

Recall of Multiple Boards

As predicted by the LT-WM theory, interfering material presented after the presentation of a position affects recall only minimally (Charness, 1976; Frey and Adesman, 1976). In the same line of research, Cooke, Atlas, Lane, and Berger (1993) show that players can recall several boards presented briefly (5 or 8 s each). Ericsson and Kintsch take these data as evidence for rapid encoding into LT-WM. However, Cooke et al. (1993), as well as Gobet and Simon (1996b), show that the percentage of correct pieces per board decreases as a function of the number of boards attempted. In addition, Gobet and Simon (1996b) show that subjects have difficulties with four and five boards.

These limits with large numbers of boards are real and seem difficult to overcome even with practice. For example, Gobet and Simon (1996b) describe the development of a master trying to increase his performance in the multiple board experiment by using a mnemonic technique to help him distinguish and store the boards. In the course of his 150-session practice, the master attempted to recall as many as 10 boards a few times but showed a plateau at around 8 boards. How does the LT-WM theory account for these results? Since Ericsson and Staszewski (1989) propose that a single retrieval structure is used by chess players, this concept cannot be used to explain the recall of multiple boards (except for the last board presented). Then, the performance is made possible through the other mechanism provided by LT-WM—LTM elaborations. But can this mechanism explain the drop in performance with four and five boards? Again, the LT-WM theory is not specific enough to answer this question.

LT-WM Theory and Chess: Summary of Evidence

In summary, the evidence adduced in favour of the application of the LT-WM theory to chess is much weaker than proposed by Ericsson and Kintsch (1995). On the one hand, the proposed hierarchical retrieval structure receives little empirical support. Contrary to the predictions of the theory, random positions with long presentation time are not better recalled than proposed by the chunking theory, manipulation of random boards is very difficult when the moves are illegal, and presentation of pieces in a scrambled order impairs recall. On the other hand, encoding through LTM elaborations seems slower than proposed by the theory, as evidenced by the multiple board recall task, where there is a drop in performance with four and five boards. Results with

concurrent and posterior interference tasks are compatible with the theory, as with most other theories. Thus, the strongest support for the LT-WM theory is offered by the ability of chess masters to rapidly access the information of individual squares and in their ability to overcome the effect of posterior interfering tasks.

Retrieval Structures in Other Domains of Expert Memory

As mentioned above, the evidence for retrieval structures is convincing for domains such as the digit span task and memory for menu orders, where there is a conscious intent to improve one's memory through strategies. Results from mental calculation and abacus calculation are compatible with the presence of a retrieval structure, but are also compatible with other theoretical explanations, such as the use of the visuo-spatial sketchpad (Baddeley, 1990; Hatano & Osawa, 1983) or the automatisation of procedures (Anderson, 1983). Contrary to the claims of Kintsch and Ericsson, I do not believe that medical expertise offers any support for retrieval structures. In several studies, an inverse U-shaped relation has been found between expertise and recall performance, intermediate-level subjects obtaining better results than experts (e.g., Schmidt & Boshuizen, 1993). Several explanations can be offered to account for this "anomaly": experts spend less time studying the material because they reach a conclusion faster; experts are used to having the information available (cf. Patel, Kaufman & Magder, 1996), and may develop the habit of not overloading their memory; or experts are more selective and focus on abstract information (the explanation favoured by Ericsson and Kintsch).[6] However, it still seems fair to say that the available empirical evidence does not offer any special corroboration for the LT-WM theory.

Indeed, it actually seems to refute two of its predictions. First, Ericsson and Kintsch (1995, p. 238) take as strong evidence for retrieval structures the fact that experts in medicine can recall pieces of information even if they are presented in a scrambled order. However, Vicente and Wang (1998) have argued that an experiment carried out by Coughlin and Patel (1987) contradicts this specific prediction (as noted above, this prediction is also incorrect with chess experts). In this experiment, dictating symptoms in a scrambled order negatively affected medical experts' recall performance when the temporal order of appearance of these symptoms was relevant for a correct diagnosis. Second, Ericsson and Kintsch (1995, p. 236-237) note that "it is reasonable to assume that a critical function of LT-WM in medical experts is not only to attain the correct diagnosis but to provide working memory support for reasoning about and evaluation of diagnostic alternatives." If this is true, medical experts should recall not only abstract information about a case, but also concrete details, as these may play an important role in discriminating between alternatives. They should therefore show better recall than intermediate-level subjects, which is not the case. I will argue below that experts do not possess the storage capacities proposed by the LT-WM theory (both retrieval structure and rapid elaboration of LTM schemas) and may encode information only after they have evoked a particular schema through pattern recognition.

In summary, the evidence for retrieval structures is good for the digit-span task and for memory for menu orders, is reasonable with mental calculation and mental abacus calculation, but it is weak for medical expertise.

An Alternative to LT-WM: Modelling Expert Behaviour

At about the same time as the LT-WM paper appeared, two other theories of expertise were published, EPAM-IV (Richman et al., 1995) and the template theory (Gobet & Simon, 1996b), which both take EPAM (Feigenbaum & Simon, 1984) as their theoretical framework. It is of considerable theoretical interest to compare LT-WM with these theories, as both deal with empirical phenomena that Ericsson and Kintsch discuss at length (digit-span case for EPAM-IV, and chess for the template theory) and both use the idea of retrieval structure. In addition, these theories are implemented as formal computer programs, which may be contrasted with Ericsson and Kintsch's informal theory.[7]

A Simulation of the Digit-Span Task Using EPAM-IV

Inspired by the insights and analyses of Chase and Ericsson (1992), Ericsson and Staszewski (1989) and Staszewski (1990), Richman, Staszewski and Simon (1995) have recently shown that the idea of retrieval structure, combined with the EPAM theory of perception and memory, offers a sufficient explanation of SF and DD's extraordinary performances in the digit-span task. Expanding and modifying earlier versions of EPAM (e.g., Feigenbaum & Simon, 1984), Richman et al. have developed a computer model that simulates DD's behaviour accurately, both at an aggregate and at a detailed level, and both qualitatively and quantitatively. The model has three main components: an STM, an LTM, and a discrimination net, which allows nodes in LTM to be accessed. Short-term memory includes specialised auditory and visual subcomponents, with iconic memories for each. LTM includes a semantic and a procedural component. Retrieval structures—which are specifically considered as deliberately acquired (p. 306)—are considered as schemas in semantic LTM. Richman et al. describe various cognitive

mechanisms in great detail, including mechanisms for learning chunks, for constructing retrieval structures, for building connections between slots of the retrieval structure and information held in semantic LTM, and for associatively searching semantic LTM. They also explicitly make clear what aspects of the theory belong to a fixed architecture, and what aspects can be changed by learning or by deliberate strategies. Finally, time parameters are associated with all processes. For example, it takes 10 ms to traverse one node of the discrimination net, and it takes 200 ms to attach information to the retrieval structure.

Richman et al.'s theory shares several features with Ericsson and Kintsch's (1995) account. Most of these features relate to the role of LTM knowledge, acquired through practice and deliberate study, in the extraordinary performance of some experts in the digit-span task. This knowledge includes retrieval structures, semantic categories (e.g., ages, running times), and numerical pattern codes (e.g., symmetry like 36-63). Their theory differs however in several ways from Ericsson and Kintsch's. In EPAM: (a) the retrieval structure contains slots, where values can be rapidly stored; (b) individual digits can be stored at the terminal nodes of the retrieval structure, and groups of digits at higher levels; (c) the retrieval structure is a schema in semantic memory; (d) approximate time parameters are given for each cognitive process, and decay parameters are estimated; and (e) times for basic cognitive processes are assumed to be invariant across skill levels. By contrast, in the LT-WM account: (a) a mechanism based on retrieval cues is used to explain how the retrieval structure permits a rapid encoding; (b) only groups of digits can be stored in the retrieval structure; (c) the retrieval structure offers encoding mechanisms different from those offered by

schemas; (d) no time or decay parameters are given; and (e) a speed-up in basic cognitive processes, such as encoding and retrieval, is postulated. Finally, the computer implementation of the EPAM theory allows one to explore complex interactions between various processes, which is not possible in the LT-WM approach. In general, Richman et al. show that a detailed theory of cognitive processes can be implemented as a computer program, which allows for a high level of precision and makes it possible to rigorously test the fit of the theory to the data. In spite of these differences, it is fair to say that Ericsson and Kintsch (1995), extending Chase and Ericsson (1982) skilled memory theory, have given an account of expertise in the digit-span task that qualitatively captures the main empirical phenomena.

Retrieval Structures and Templates: The Case of Chess Expertise

As discussed above, the empirical data on chess skilled memory only weakly support the hypothesis of a 64-square hierarchical retrieval structure and even contradict some of the predictions of the LT-WM theory, mainly because the proposed mechanisms (retrieval structure and rapid generation of LTM structures) are too powerful. In addition, the lack of specificity of LT-WM account of chess expertise makes it difficult to evaluate. Since Ericsson and Kintsch (1995) claim that data from chess research provide some of the most compelling evidence for LT-WM, this implies serious limitations in the generality of their theory. Although the interference and multiple board experiments suggest that the general message of Ericsson and Kintsch—LTM encoding is faster than was supposed in earlier theories—is valid, it appears that chess masters use mechanisms different from that postulated by Ericsson and Kintsch (1995).

Why does the LT-WM theory, which convincingly explains mnemonists' behaviour, have so many difficulties with data from expert chess memory? It is not because chess is an idiosyncratic task removed from normal human cognition: the main features of chess expertise (selective search, memory for meaningful material in the domain of expertise, importance of pattern recognition) are generalizable to quite a few other domains (Charness, 1992). Rather, these difficulties seem to come from inappropriately generalising theoretical constructs that are specific to domains where explicit mnemonic strategies play an important role. More specifically, the LT-WM approach postulates for many domains what I will call a <u>generic</u> retrieval structure (cf. Figure 4 of Ericsson & Kintsch, 1995), that is a structure that does not depend on the specifics of the material presented (assuming that this material belongs to the domain of expertise); while this structure allows better storage when the input can be coded as high-level representations such as schemas, it can also be used when the material is less well structured or not structured at all. A key feature of a generic retrieval structure is therefore that it can be used even with random material. Moreover, in at least some task domains, such as mental abacus calculation, this structure permits information to be encoded through retrieval cues even in the absence of the second encoding mechanism provided by the LT-WM theory—integration through schemas already stored in LTM.

A comparison of the LT-WM theory and the template theory (Gobet & Simon, 1996b; 1998; in press) will clarify what are the main theoretical weaknesses, in semantically rich domains, of the type of general-purpose retrieval structures proposed by Ericsson and Kintsch. The template theory, which is implemented as a computer program, is a revision of the chunking

theory (Chase & Simon, 1973) and, like Richman et al.'s (1995) model, it uses the EPAM architecture, including its time parameters (there are some minor differences that need not concern us here). It proposes that expertise is due to: (a) a large database of chunks, indexed by a discrimination net; (b) a large semantic memory, including both schemas and productions; and (c) associations between the (perceptual) chunks and the semantic memory, as well as associations within semantic memory. In addition, the theory proposes that some chunks that recur often during practice and study evolve into more complex data structures (templates), which allow information to be rapidly encoded into slots. For example, in a given class of chess positions, a slot may be created for a square that plays an important role and that can be occupied by different types of pieces or be empty at different times. Templates offer both core information that cannot be changed (as with the chunks of the chunking theory) and slots allowing rapid updating of information. The theory proposes that different templates will be retrieved for different classes of positions and that, because of their construction based on the recurrence of certain patterns of pieces, templates are unlikely to be recognised in random positions. Finally, the template theory assumes that visual STM is limited to about three items (cf. Zhang & Simon, 1985). The template theory accounts for most empirical data from chess memory research (Gobet, 1998a), and is also compatible with data on problem solving (Gobet, 1997).

While the template theory has much in common with the LT-WM theory, in particular the assumption that experts are able to encode material into LTM rapidly and that this ability is limited to their domain of expertise, there are two basic differences in their use of retrieval structures, one related

to the retrieval availability, the other to the specificity of encoding: (a) Templates are specific to a class of stimuli and can be used only in cases where they have been accessed through perceptual cues, while the retrieval structure proposed by Ericsson and Kintsch for chess is generic; and (b) Template slots are created only for certain aspects of the stimuli (in chess, certain squares or pieces, depending on the class of positions), while the LT-WM theory assumes that the retrieval cues of the retrieval structures apply equally to all stimuli (the 64 "squares" of the chess retrieval structure are usable for any type of position, as well as the other (unspecified) levels of the hierarchical retrieval structure). These two differences explain why one theory accounts successfully for most of the results reviewed here, and why the other fails. In particular, three sets of results follow directly from the template theory: first, the necessity of accessing chunks and templates in LTM through recognition explains that recall is more difficult with random positions and positions presented with non-natural groupings of pieces; second, rapid access by recognition through the discrimination net accounts for the high performance with short presentation times; and third, strong players can recall several boards presented rapidly because they can use one template per position if the positions are typical enough; the limit of visual STM and the slow LTM storage time (about 8 s to create a chunk) impose a limit of about 4 items, which accounts for masters' difficulty in remembering more than four boards. (As mentioned above, the second interpretation of LT-WM accounts for the first sets of results as well. However, all the explanatory power is given by recognition of chunks—patterns or schemas—and not by the use of a retrieval structure or elaboration of LTM structures).

Types of Retrieval Structures

What Ericsson and Kintsch refer to as retrieval structures consist of (at least) three different sets of entities. First, there are the generic retrieval structures used in the digit-span task, the restaurant menu task, and, for that matter, in most mnemonics such as the method of loci. In these tasks, experts use arbitrary structures, developed with the deliberate goal of improving their memory (cf. the modelling of the digit-span task by Richman et al., 1995). Tasks like mental abacus calculation and mental calculation may also fit into this category, with the qualification that the retrieval structures are not arbitrary but are shaped by the demands of the tasks (e.g., the necessity of computing intermediate products) and with the complication that encoding into LTM does not seem to allow as good retrieval as in the previous tasks (cf. the moderate to poor incidental recall), perhaps because there is a deliberate intention to encode cues that are useful only for a limited amount of time. Finally, as argued above, chess experts' retrieval structures were incorrectly identified by Ericsson and Kintsch (1995) as being generic.

Second, there is the episodic text structure, which is built up rapidly with the comprehension of a text, but which becomes more rapidly inaccessible than the first type of structures. Third, domain knowledge—that is, patterns and schemas—is also taken to provide retrieval structures in text comprehension. These three types of retrieval structures have clearly different properties (stability, relation with schemas, generality of use), and it is unfortunate that Ericsson and Kintsch have conflated them into a single construct.

Templates, which may be construed as specific, schematic retrieval structures within the Ericsson and Kintsch framework, can be used only when cues in the environment signal their appropriateness, and their usefulness

cannot be anticipated beforehand. Tasks like chess, medical diagnosis, and to some extent text comprehension seem to fit into this category. The difference between generic and specific retrieval structures is by no means trivial: within the template theory, specific retrieval structures are included in LTM schemas, while Ericsson and Kintsch make it clear that their generic retrieval structures are distinct from, although interconnected with, LTM schemas (e.g., their Figure 4, and the discussion surrounding it). (See also Gobet & Simon, 1996b, p. 36-38, and Richman et al., 1995, p. 306, for a comparison of generic and specific retrieval structures.)

Conclusion

The concept of retrieval structure, in the strict sense of a generic structure, does offer a theoretically plausible explanation in domains where memory for order is important, where there is a conscious effort to both construct and use a memory structure under strategic control, and where the input is encoded serially. It does not fit with the empirical data in the domain of chess and medical expertise, where specific structures seem to be used. Establishing what categories of experts use such generic retrieval structures is an empirical question, although a cursory examination of the task environments in which experts operate indicate that information is often not encoded into an abstract, generic structure but is integrated into semantic networks already present in LTM. While Ericsson and Kintsch do consider the latter type of encoding, they also give an unduly important role to generic retrieval structures.

In summary, the empirical support for the LT-WM theory is much weaker than proposed by Ericsson and Kintsch. With some data (in particular in chess and medical expertise), the predictions of the theory are directly

refuted. With others, the theory is too vague to generate predictions without numerous auxiliary assumptions. Finally, the theory is not internally consistent, the key concept of retrieval structure referring to entities having different psychological properties. Ericsson and Kintsch's review of empirical results from the literature on expertise and text comprehension represents a welcome challenge to theories of working memory, and this is an important contribution of their paper. However, the LT-WM theory—their own response to this challenge—fails, because the evidence provided for LT-WM mechanisms and structures is rather weak.

Implicit in my discussion of LT-WM is the assumption that psychological theories should be stated precisely and unambiguously, if possible as mathematical or computational models. While Ericsson and Kintsch mention in their conclusion (1995, p. 240) the need of implementing LT-WM as a processing model, they also think that an informal account can be given that makes testable predictions. In this respect, my position is less optimistic: given the complexity and the number of mechanisms involved in a theory such as LT-WM—encodings into retrieval structures, elaborative encodings, strategies, proactive and retroactive interference, amount of expert knowledge used, interaction with the task demands—I believe it hopeless to try to derive predictions and explanations without a process model implemented as a computer program. Note that LT-WM is in a better situation with text comprehension, where subsets of the theory have been implemented as computer programs (see Kintsch, 1998).

This bring us to an important methodological question: How to compare informal theories covering a large number of domains with little precision with formal theories covering a small number of domains with high

precision? There are at least three reasons, in addition to those mentioned above, why formal, quantitative theories should be preferred: (a) their fit to empirical data can be estimated, which may indicate places where the theory can be improved; (b) their use of quantitative parameters allow quantitative predictions, which are always stronger than qualitative predictions; and (c) their formal structure makes it harder to use ad hoc assumptions, which in turn keeps the number of degrees of freedom of the theory under control. (For a detailed discussion of these points, see Grant, 1962; Gregg & Simon, 1967; Meehl, 1967; and Popper, 1959).

As a conceptual framework, loosely combining mechanisms and structures, but summarising a large body of empirical data and hinting at where future research should be carried out, there is little to criticise in LT-WM. But, as a theory—and by comparing LT-WM to other current theories of memory, Ericsson and Kintsch invite us to consider it as such—LT-WM is disappointing in its current state: key concepts are not defined, many mechanisms, structures and parameters are left unspecified, and several empirical findings are at variance with the theory. Now, it is true that Ericsson and Kintsch's paper presents the first version of their theory and that several of these weaknesses will hopefully be removed in future elaborations. My expectation is that, as the theory is modified and improved, the need for a formal, information-processing implementation will become more apparent.

Reference list

Anderson, J.R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Anderson, J.R., & Lebière, C. (1998). *The atomic components of thought*. Mahwah, NJ: LEA.

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

Binet, A. (1966). Mnemonic virtuosity: A study of chess players. *Genetic Psychology Monographs, 74*, 127-162. (Translated from the Revue des Deux Mondes (1893), 117, 826-859).

Charness, N. (1976). Memory for chess positions: Resistance to interference. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 641-653.

Charness, N. (1992). The impact of chess research on cognitive science. *Psychological Research, 54*, 4-9.

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. (1973). 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. and Berger, R.C. (1993). Role of high-level knowledge in memory for chess positions. *American Journal of Psychology, 106*, 321-351.

Coughlin, L.D., & Patel, V.L. (1987). Processing of critical information by physicians and medical students. *Journal of Medical Education, 62,* 818-828.

de Groot, A.D. (1978). *Thought and choice in chess*. The Hague: Mouton Publishers. First Dutch edition in 1946.

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

Delaney, P.F., & Ericsson, K.A. (1998). Working memory and expert performance. In R.H. Logie & K.J. Gilhooly (Eds.), *Working memory and thinking*. London: Psychology Press.

Ericsson, K.A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211-245.

Ericsson, K.A., & Oliver, W. (1984, November). *Skilled memory in blindfolded chess.* Paper presented at the annual meeting of the Psychonomic Society, San Antonio, TX.

Ericsson, K.A., & Staszewski, J.J. (1989). Skilled memory and expertise: Mechanisms of exceptional performance. In D.K. Klahr & K. Kotovski (Eds.), *Complex information processing: The impact of Herbert A. Simon.* Hillsdale, NJ: Erlbaum.

Feigenbaum, E.A., & Simon, H.A. (1962). A theory of the serial position effect. *British Journal of Psychology, 53*, 307-320.

Feigenbaum, E.A., & Simon, H.A. (1984). EPAM-like models of recognition and learning. *Cognitive Science*, *8*, 305-336.

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 [The memories of a chess player].* Fribourg: Editions Universitaires.

Gobet, F. (1997). Roles of pattern recognition and search in expert problem solving. *Thinking and Reasoning, 3,* 291-313.

Gobet, F. (1998a). Expert memory: Comparison of four theories. *Cognition, 66,* 115-152.

Gobet, F. (1998b). Memory for the meaningless: How chunks help. *Proceedings of the 20th Meeting of the Cognitive Science Society* (pp. 398-403). Mahwah, NJ: Erlbaum.

Gobet, F., & Simon, H.A. (1996a). Recall of rapidly presented random chess positions is a function of skill. *Psychonomic Bulletin & Review, 3,* 159-163.

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. (1998). Expert chess memory: Revisiting the chunking hypothesis. *Memory,* 6, 225-255.

Gobet, F., & Simon, H.A. (in press). Five seconds or sixty? Presentation time in expert memory. *Cognitive Science*.

Grant, D. A. (1962). Testing the null hypothesis and the stategy and tactic of investigating theoretical models. *Psychological Review, 69,* 54-61.

Gregg, L.W., & Simon, H.A. (1967). Process models and stochastic theories of simple concept formation. Journal of Mathematical Psychology, 4, 246-276.

Hatano, G., & Osawa, K. (1983). Digit memory of grand experts in abacus-derived mental calculation. *Cognition, 15,* 95-110.

Holding, D.H. (1985). *The psychology of chess skill.* Hillsdale, NJ: Erlbaum.

Kinstch, W. (1998). *Comprehension. A paradigm for cognition.* Cambridge, UK: Cambridge University Press.

Meehl, P.E. (1967). Theory testing in psychology and physics: A methodological paradox. *Philosophy of Science*, 103-115.

Newell, A. (1990). *Unified Theories of Cognition*. Cambridge, MA: Harvard University Press.

Newell, A., & Simon, H.A. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.

Patel, V.L., Kaufman, D.R., & Magder, S.A. (1996). The acquisition of medical expertise in complex dynamic environments. In K.A. Ericsson (Ed.), *The road to excellence*. Mahwah, NJ: Lawrence Erlbaum.

Popper, K.R. (1959). *The logic of scientific discovery.* London: Hutchinson.

Richman, H.B., Staszewski, J., & Simon, H.A. (1995). Simulation of expert memory with EPAM IV. *Psychological Review, 102*, 305-330.

Robbins, T.W., Anderson, E.J., Barker, D.R., Bradley, A.C., Fearnyhough, C. Henson, R., Hudson, S.R., & Baddeley, A.D. (1996). Working memory in chess. *Memory & Cognition, 24*, 83-93.

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. (1992). Visuospatial and articulatory interference in chess players' information intake. *Applied Cognitive Psychology*, *6*, 77-89.

Schmidt, H.G., & Boshuizen, H.P.A. (1993). On the origin of intermediate effects in clinical case recall. *Memory & Cognition, 21,* 338-351.

Staszewski, J.J. (1990). Exceptional memory : The influence of practice and knowledge on the development of elaborative encoding strategies. In F.E. Weinert & W. Schneider (Eds.), *Interactions among aptitudes, strategies, and knowledge in cognitive performance*, (pp. 252-285). New York: Springer.

Vicente, K. J., & Wang, J. H. (1998). An ecological theory of expertise effects in memory recall. *Psychological Review, 105*, 33-57.

Zhang, G., & Simon, H.A. (1985). STM capacity for Chinese words and idioms: Chunking and acoustical loop hypothesis. *Memory and Cognition*, *13*, 193-201.

Footnotes

[1]Basic storage and retrieval operations seem to vary as a function of expertise in the LT-WM theory. However, without an indication of at least the order of magnitude of time and decay parameters and of their rate of change with expertise, it is impossible to derive any quantitative prediction from the theory.

[2] This assumes a serial encoding within the retrieval structure. This seems in agreement with Ericsson and Kintsch, who state that experts use retrieval structures similar to those of SF and do not make any mention of parallel encoding.

[3] One has only to assume that it takes about 8 s to create a new chunk (Newell & Simon, 1972), that it takes about 2 s to add information to a chunk (Simon, 1976), and that the last chunks (containing possibly only one piece) attended to are stored in STM at the end of the presentation. This explanation works both for visual and auditory presentation (see Gobet, 1998, and Gobet & Simon, in press, for computer simulations of experiments with visual presentation). Thus, contrary to Ericsson and Kintsch's (1995, p. 237) claims, theories of chess skill based on visual recognition of chunks can account for experiments where the pieces are dictated one at a time.

[4] Cf. Delaney and Ericsson (1998, p. 109): "The maintenance of encoded relations between pieces and locations in a random position therefore appear to depend critically on the retrieval structure, which can only uniquely index a single chessboard arrangement at a time."

[5] Actually, it should be easier to encode them into new distinct structures in LTM, because random positions differ radically from the positions learnt by chess players through experience and, therefore, should suffer less mutual interference.

[6] As noted by an anonymous reviewer, age is confounded with expertise in almost every medical expertise study. Given that memory in recall tasks is age sensitive, this confound makes the interpretation of these studies very difficult indeed.

[7] The emphasis here is in comparing LT-WM with these two theories. For a detailed account of EPAM-IV, including the differences with previous versions of EPAM and a discussion of the simulations, the reader is referred to Richman et al. (1995). Earlier versions of CHREST (Chunk Hierarchy and REtrieval STructure), the computer implementation of the template theory, are discussed in Gobet (1993) and in De Groot & Gobet (1996); the current version is discussed in Gobet (1998) and in Gobet and Simon (in press).