



How chunks, long-term working memory and templates offer a cognitive explanation for neuroimaging data on expertise acquisition: A two-stage framework

Alessandro Guida^{a,b,*}, Fernand Gobet^c, Hubert Tardieu^d, Serge Nicolas^a

^a Institut de Psychologie, Université Paris Descartes, Boulogne Billancourt, France

^b CRPCC, Université Rennes 2, Rennes, France

^c Centre for Cognition and Neuroimaging, Brunel University, Uxbridge, United Kingdom

^d Institut de Psychologie, Université Paris Descartes, EPHE, Boulogne Billancourt, France

ARTICLE INFO

Article history:

Accepted 19 January 2012

Available online 29 April 2012

Keywords:

Expertise

Chunks

Long-term working memory

Template theory

Brain functional reorganization

Working memory

ABSTRACT

Our review of research on PET and fMRI neuroimaging of experts and expertise acquisition reveals two apparently discordant patterns in working-memory-related tasks. When experts are involved, studies show activations in brain regions typically activated during long-term memory tasks that are not observed with novices, a result that is compatible with functional brain reorganization. By contrast, when involving novices and training programs, studies show a decrease in brain regions typically activated during working memory tasks, with no functional reorganization. We suggest that the latter result is a consequence of practice periods that do not allow important structures to be completely acquired: knowledge structures (i.e., Ericsson and Kintsch's retrieval structures; Gobet and Simon's templates) and in a lesser way, chunks. These structures allow individuals to improve performance on working-memory tasks, by enabling them to use part of long-term memory as working memory, causing a cerebral functional reorganization. Our hypothesis is that the two brain activation patterns observed in the literature are not discordant, but involve the same process of expertise acquisition in two stages: from decreased activation to brain functional reorganization. The dynamic of these two physiological stages depend on the two above-mentioned psychological constructs: chunks and knowledge structures.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

One of the fundamental questions about working memory (WM) concerns its limit. Miller (1956) famously proposed that the amount of information that can be kept in mind at one time is about seven chunks or meaningful units of information. More recently, Ericsson and Kintsch (1995), with their long-term working memory theory (LT-WMT), and Gobet and Simon (1996a), with their template theory (TT), have proposed that, in the case of expertise, part of long-term memory (LTM) can be used during WM tasks in order to circumvent the limit imposed by the magical number 7 (Miller, 1956); this would explain, for example, the performance of experts who are able to recall more than 100 digits (e.g., Chase & Ericsson, 1981). Recent developments in brain imaging via functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) have provided partial support for this idea. Two patterns of results can be observed during WM-related tasks. When the experiments involve expert individuals, these do

indeed tend to show an activation of brain regions typically activated during LTM tasks (LTM areas hereafter). This pattern of results could correspond to a cerebral functional reorganization related to the acquisition of expertise, where functional reorganization is seen as the recruitment of new activation areas and a shift in the cognitive process underlying task performances (Poldrack, 2000), which in this case means that resources previously (before expertise acquisition) allocated to WM are later (after expertise acquisition) allocated to LTM.

By contrast, when the experiments involve novices who undergo extended practice with WM-related tasks, the results tend to show a decrease in activation of brain regions typically activated during WM tasks (WM areas hereafter), with no cerebral functional reorganization. The latter result could be a consequence of practice periods that do not allow important cognitive structures to be completely acquired, mainly knowledge structures¹ (i.e., retrieval structures, Ericsson & Kintsch, 1995; templates, Gobet & Simon, 1996a) and in a less important way chunks (Chase & Simon, 1973a).

In this paper, after presenting chunking theory and two more recent theories of expertise (LT-WMT and TT), we will first

* Corresponding author at: Centre de Recherche en Psychologie, Cognition et Communication, Université Rennes 2 – Haute Bretagne, Place du Recteur Henri Le Moal, CS 24 307, Bâtiment S, 35 043 Rennes Cedex, France.

E-mail addresses: alessandro.guida@univ-rennes2.fr, alessandro.guida.psychology@gmail.com (A. Guida).

¹ We use “knowledge structures” as an umbrella term for retrieval structures (Ericsson & Kintsch, 1995) and templates (Gobet & Simon, 1996a).

establish using the Bayes Factor that it is possible to separate episodic long-term memory activation from working memory activation and then review the studies that report the two different brain activation patterns mentioned above. We will then propose an explanation that makes these results more coherent by suggesting that the two brain patterns are two stages of the same process occurring during expertise acquisition. We will conclude by discussing the psychological constructs – chunks and knowledge structures – that underpin expertise in this two-stage view, linking them with physiological processes.

2. Chunking theory

One important element for understanding WM limits is chunking. This mechanism is important not only in understanding standard cognitive performance, but also in explaining the differences between novices and experts. The chunking mechanism was initially described by de Groot (1946/1978) and Miller (1956), and then theorized by Chase and Simon (1973a). A current definition is given by Gobet et al. (2001, p. 236): a chunk refers to "... a collection of elements having strong associations with one another, but weak associations with elements within other chunks." To explain chunks, Miller (1956) and subsequently Cowan (2001) used the same example. When the letters "fbiiibm" are presented, if one knows the acronyms "FBI" and "IBM," then it is possible to simplify the information by forming two chunks² ("FBI" and "IBM") in WM. Because these familiar patterns exist in LTM – FBI is the Federal Bureau of Investigation and IBM is a well-known computer company – the letters "f," "b," "i," "i," "b" and "m" can be encoded as two elements in WM instead of six.

Chunking theory provides an explanation for the superiority of experts over novices. For example, in the field of chess, chunking has been used to explain how chess experts are able to recall more chess pieces on a board than novices (Chase & Simon, 1973b; de Groot, 1946/1978). Thanks to their greater knowledge of chess positions – in terms of chunks in LTM – the experts are able to encode the presented chess positions in fewer chunks in WM, thereby gaining storage space in WM. For example, a chess master can encode 15 pieces presented on a board as one chunk (Gobet & Clarkson, 2004). This process is statistically less likely for novice players since they have less knowledge of chess positions and therefore possess fewer chunks in LTM.

Chase and Simon (1973a) confirmed de Groot's (1946/1978) results, but also found that when the chess pieces to be recalled were placed in random locations, the superiority of experts over novices disappeared. They argued that experts could not use their LTM chunks to encode the random positions, since these positions did not contain any of the chunks the experts had in LTM. Thus, the positions were mostly as new to them as to the novices. The experts' advantage was significant only when they could actually use their knowledge. With their MAPP (*Memory Aided Pattern Perceiver*) computer simulation, Simon and Gilmartin (1973) estimated that the number of chunks in LTM to reach a master level ranged from 10,000 to 100,000.

2.1. Weaknesses of chunking theory

Since 1973, new data have led to a revision of some of the results presented above. For example, Gobet and Simon (1996b, p. 159) established, through a review of chess experiments where

² The term "chunk" can be employed to mean both "the elements that are in WM" and "the elements that are in LTM." To distinguish between these two meanings, some researchers prefer to use the term "pointers" to refer to the content of WM that is linked (that points) to LTM knowledge. When the distinction is required, we will use the terms "chunks in WM" and "chunks in LTM."

random positions served as control material, that "strong players generally maintain some superiority over weak players even with random positions, although the relative difference between skill levels is much smaller than with game positions." Using the CHREST (*Chunk Hierarchy and REtrieval Structures*) computer simulation, the same authors estimated that the number of chunks required in LTM to reach a master level was about 300,000 (Gobet & Simon, 2000).

The new studies have helped clarify the chunking mechanisms but have also cast doubt on one crucial point of the theory. Chunking theory, as developed by Chase and Simon (1973a), hypothesizes that information, once encoded as chunks (also called pointers), is stored in WM. But, as Gobet (2000a) has noted, (a) it is difficult to understand how very complex tasks can be solved by juggling chunks in a limited space like WM, and (b) the information encoded as chunks (or at least part of these chunks) may be stored in LTM.

Concerning this last point, Charness showed as early as in 1976 that an interfering task between the presentation of a chess position and the test phase reduced the recall performance of experts only marginally. If the chunks were stored in WM, as theorized by Chase and Simon (1973a), then at least some of them should have been erased by the interfering task. But, since this did not occur, Charness (1976) suggested that at least some chunks were rapidly transferred to LTM. In the same year, Frey and Adesman (1976) observed a similar result; more recently, Cooke, Atlas, Lane, and Berger's (1993) study confirmed it. In addition, similar results have been found in text comprehension using an interfering task (Glanzer, Dorfman, & Kaplan, 1981; Glanzer, Fisher, & Dorfman, 1984). The absence of the effect of the interfering task has been interpreted as being due to LTM storage (Ericsson & Kintsch, 1995; Gobet & Simon, 1996a; Kintsch, 1998). The idea that part of LTM can be used to store information during WM tasks also explains the performance of experts who are able to recall more than 100 briefly presented digits (Chase & Ericsson, 1981; for a complete review, see Ericsson & Kintsch, 1995).

To sum up, it seems that processes supposed to occur solely in WM also involve, in the case of expertise, LTM storage. As emphasized by Gobet (2000a, p. 552), the information is "transferred in LTM more rapidly than suggested by the chunking theory." It is following this idea that Ericsson and Kintsch (1995) and Gobet and Simon (1996a) proposed their respective theories.

3. Long-term working memory and template theory: a same core idea

LT-WMT and TT revolve around the same fundamental core idea: with expertise, part of LTM can be used as WM, thus expanding an individual's memory storage and processing capacities. In both theories, this is possible only when knowledge structures have been built. These structures have been called "retrieval structures" in Ericsson and Kintsch's (1995) theory and "templates" in Gobet and Simon's theory (1996a). We briefly introduce the principal features of these two structures.

3.1. Template theory

Templates (Gobet & Simon, 1996a) can be seen as schemas with slots that can be filled with variable information, including chunks. An example would be the schema of a train station that may (or may not) be filled with a train. Chunks and templates are both typical configurations, but, while chunks are simple patterns, templates are configurations that can be filled with additional information, such as pieces in chess. In other words, templates are high-level patterns (prototypes) that can change their aspect slightly – via

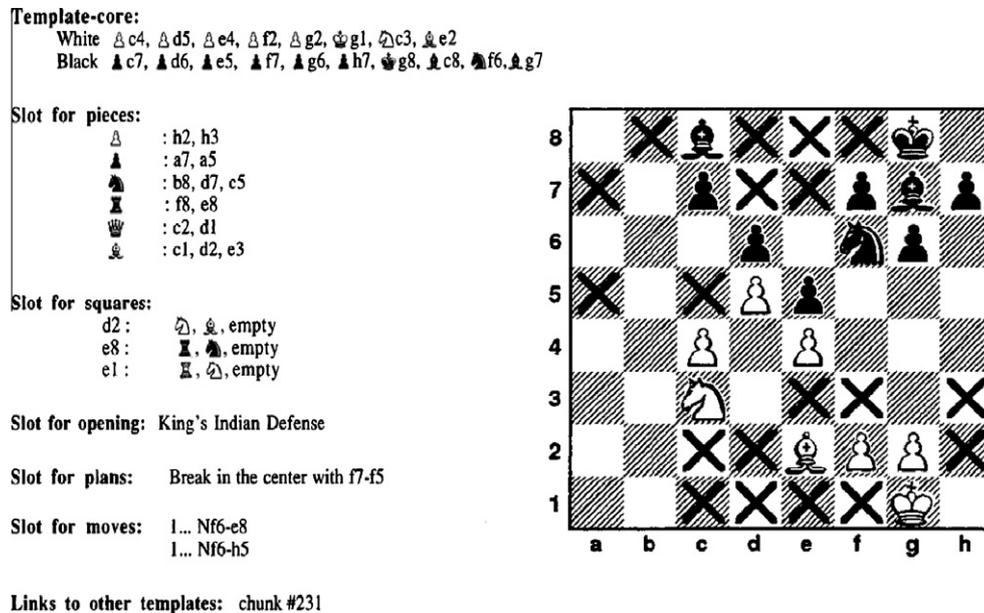


Fig. 1. Example of a template from Gobet and Simon (1996a). On the left, the attributes of the template: the non-variable part of a template (the template-core) and the variable part of a template (slots for pieces, squares, openings, plans, moves, and links to other templates). On the right, the diagrammatic representation of the same template: pieces on the board indicate the core pieces in the template and crosses indicate values contained in piece or square slots.

the filling of slots – but that are still categorized in the same way. Thus, the same template can be used by an expert to encode different chess positions, assuming that they belong to the same category.

Gobet and Simon (1996a) presented an example of a template (see Fig. 1). If a complete novice looks at a typical position from the King's Indian Defense – of which the chess board in Fig. 1 depicts some of the key pieces – the pattern will not really be meaningful. However, if a fairly strong chess player looks at such a position, then two templates might be perceived: a first template for the white pieces and a second template for the black pieces. If an even stronger player – say a grandmaster – looks at the position, then the two templates are likely to be perceived as a single template. A template has a core (the non-variable part of the template) and slots (the variable part). On the chessboard in Fig. 1, the pieces on the board indicate the core, while the crosses indicate the positions where chess pieces can be added without changing the template. This means that a player will categorize different but related positions with the same template. For example, a white pawn can be added on h2 or h3, and a black knight can be added on d7. A template can also be linked to other types of information: openings, moves, plans, and other templates. TT has since been partly implemented with CHREST. Based on simulations with this model, Gobet and Simon (2000) have estimated that it takes 250 ms to fill a template slot.

3.2. Long-term working memory theory

LT-WMT (Ericsson & Kintsch, 1995) differs from TT in that the storage of information by an expert is not considered as the particularization of a schema, but as an association between the encoded information and a set of “retrieval” LTM cues. In order to retrieve the encoded information, the expert must reinstate the encoding conditions by using the same set of retrieval cues. Long-Term Working Memory becomes available (but restricted to the field of expertise), when a set of cues becomes a stable structure in LTM: a retrieval structure. Ericsson and Kintsch's theory is a generalization of Chase and Ericsson's (1981) skilled memory theory and applies the same three principles: meaningful encoding, structured retrieval, and acceleration of encoding and retrieval. The first principle states that, to be easily stored in LTM, the incoming information needs to be

transformed into meaningful units. The second principle concerns the retrieval structure (see below for an example). And the last principle states that, with extensive practice, experts can speed up the encoding and retrieval of information in LTM.

Fig. 2 gives an example – inspired by S.F.'s (a runner who used his knowledge of footraces to encode and recall digits as running times) performance and verbal protocols analyzed and described by Chase and Ericsson (1981; see also Ericsson, 1985) – of a retrieval structure combined with supplementary associations called elaborated memory structure. These supplementary associations are “knowledge-based associations relating units of encoded information to each other along with patterns and schemas establishing an integrated memory representation...” (Ericsson & Kintsch, 1995, p. 221).

In Fig. 2, “146732981417” is the sequence that must be recalled. If one has a good knowledge of times for footraces (like S.F. had), the sequence can be encoded by coding the digits as running times for races (the principle of meaningful encoding). Each running time is associated with one specific cue: “first” for the first running time, “middle” for the second running time, and “last” for the third time. Therefore, “146732981417” becomes “1 min 46.7 s” linked to “first,” “3 min 29.8 s” linked to “middle,” and “1 min 41.7 s” linked to “last.” To recall the information, one must just activate the cues: first, middle and last (the principle of structured retrieval). Moreover, supplementary associations can also be created as the information is encoded following the retrieval structure. In this example, the encoded stimuli can be associated by means of knowledge of running times; hence, the first and last times can be associated since they are both 800-m times, while the last two times can be linked together since they are both very good running times. These supplementary associations vary according to the incoming information and help the individual to have “an integrated memory representation” (Ericsson & Kintsch, 1995).

3.3. Long-term working memory and template theory specificity compared to working-memory models and theories

Even though, as we have just explained, some differences exist between LT-WMT and TT (for an analysis, see Cowan, 2005; Ericsson & Kintsch, 2000; Gobet, 1998a, 2000a, 2000b), both

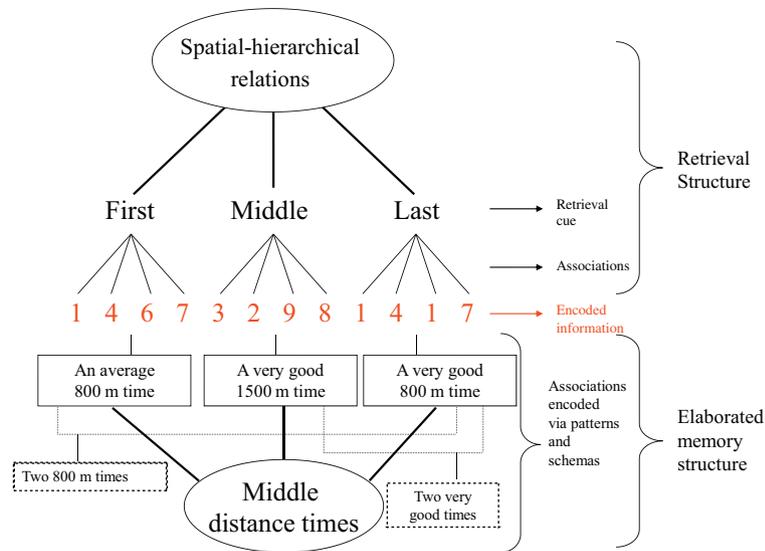


Fig. 2. Example of a retrieval structure and an elaborated memory structure, adapted from Ericsson and Kintsch (1995) and Ericsson and Delaney (1999).

theories propose that, when an individual has acquired expertise in a domain, part of LTM can be accessed rapidly and reliably and therefore used as WM. This proposal can be compared with WM models or theories that include immediate access to LTM elements that are available for processing but are not actively maintained by attention processes (e.g., Cowan, 1995; Just & Carpenter, 1992; Oberauer, 2002). But it differs from these models concerning the description and the influence of the LTM structures that are acquired with expertise: knowledge structures. This central feature makes LT-WMT and TT different from models based on the classical definition of WM, that is, “temporary storage of information that is being processed in any of a range of cognitive tasks” (Baddeley, 1986, p. 34) particularly if by “temporary storage” one intends “active maintenance” or “rehearsal.”

In the domain of WM models, the dominant point of view – known as the “resource-sharing model” (Hitch, Towse, & Hutton, 2001) – considers that WM depends on cognitive resources, resources that are supposed to be a pool shared by short-term storage and processing. This resource-sharing hypothesis has dominated WM literature at least since Daneman and Carpenter (1980). Even if other viewpoints have been advanced (e.g., Engle, Kane, & Tuholski, 1999; Hasher & Zacks, 1988; MacDonald & Christiansen, 2002; Maehara & Saito, 2007; Towse, Hitch, & Hutton, 1998, 2000; Waters & Caplan, 1996; for a review, see Miyake, 2001) that do not appeal to the idea of resource sharing, they “are still based on a core assumption, which is also at the root of the resource sharing hypothesis, that *memory items are “actively” maintained during performance on working memory span tests*” (Saito, 2006, p. 54, italics added).

LT-WMT and TT clearly assume that WM tasks do not only measure WM in the form of a temporary storage, but also as a capacity to massively retrieve information from LTM via knowledge structures. This allows LT-WMT and TT to explain data that are difficult to interpret otherwise, for instance, how skilled activities can be interrupted and resumed later with no detrimental effects on performance (e.g., Charness, 1976; Cooke et al., 1993; Frey & Adelman, 1976; Glanzer et al., 1981, 1984). If one assumes that the elements in WM can be associated to elements in LTM (i.e., knowledge structures) and can thus be rapidly transferred to LTM in order to later be retrieved, the absence of a disruption effect becomes understandable.

3.4. Physiological consequences of long-term working memory and template theory

In theory, the capacity to use LTM structures during WM tasks has also physiological consequences. In fact, *in the case of experts*, if the incoming information instead of being stored simply in WM is rapidly linked to knowledge structures – which depend on semantic memory – and if the result of the association is stored as an episode in episodic memory, then one can assume a considerable involvement by both WM and LTM brain areas during WM-related tasks. This should contrast with the involvement of almost only WM brain areas *in the case of novices*. In other words, from the point of view of both Ericsson and Kintsch’s and Gobet and Simon’s theories, a cerebral functional reorganization (Kelly & Garavan, 2005) implicating LTM areas is expected with expertise acquisition: Here, cerebral functional reorganization does not simply mean that experts do the same mental operations as novices but with different brain areas. Rather, it means that experts do the tasks differently, that is, they execute WM tasks using different mental operations based on LTM areas.

However, while behavioral data in various domains seem to corroborate the idea of LTM storage during WM-related tasks, with the involvement of both WM and LTM in the case of experts (for examples in air piloting, Sohn & Doane, 2003; in chess, Gobet & Clarkson, 2004; Gobet & Jackson, 2002; Gobet & Simon, 1996a, 1998; Saariluoma & Kalakoski, 1997; in the game of Go, Masunaga & Horn, 2000; in gymnastic floor routines, Tenenbaum, Tehan, Stewart, & Christensen, 1999; in mnemonists, Ericsson, Delaney, Weaver, & Mahadevan, 2004; in reading span, Guida, Tardieu, & Nicolas, 2009; in text comprehension, Guida & Tardieu, 2005; Hupet, Schelstraete, Demanet, & Pourtois, 2000; in verbal knowledge of football, Postal, 2004; and in written production, Kellogg, 2001), neuroimaging results in PET and fMRI studies provide only partial confirmation. In fact, in WM-related tasks, the neuroimaging literature on experts and expertise acquisition exhibits two contrasted patterns of results. While experts’ neuroimaging studies are fairly congruent with Ericsson and Kintsch’s and Gobet and Simon’s predictions of a cerebral functional reorganization implicating LTM areas, the studies involving trained novices are mostly consistent with a decrease in cerebral activity related to practice, with no cerebral functional reorganization.

Before reviewing the studies having shown these two different patterns of results, we need to put forward the assumptions that will guide us throughout this paper and some of the issues that stem from these assumptions.

4. Assumptions, limitations and physiological predictions of the study

4.1. Separating episodic long-term memory activation from working memory activation

Our main assumption is that it is possible to separate brain activations that are from episodic LTM and brain activations that are from WM. In the literature, there seems to be evidence supporting the idea that WM and episodic LTM are different mental functions underpinned by specific and different neural pathways (for reviews, see for example Baddeley, 2003; Gazzaniga, Ivry, & Mangun, 2009; Kandel, Schwartz, & Jessell, 2000; Smith & Kosslyn, 2007; Squire & Wixted, 2011). However, the main difficulty when defending this point of view is the overlap that does exist between WM and LTM structures. For example, a meta-analysis that compared regions activated during verbal LTM and short-term memory (STM) tasks clearly showed overlapping regions in frontal and parietal lobes (Cabeza, Dolcos, Graham, & Nyberg, 2002; Cabeza & Nyberg, 2000). Nevertheless, we would like to argue that there is enough evidence showing that the medial temporal lobe is the key element to separate episodic LTM from STM/WM. For over 50 years, data from neuropsychology have clearly linked the medial temporal with LTM and not with STM/WM (e.g., Cave & Squire, 1992; Müller & Knight, 2006; Scoville & Milner, 1957; Shallice & Warrington, 1970). A large number of neuroimaging studies have also confirmed this, exhibiting a link between episodic memory and the medial temporal lobe (for reviews, Eichenbaum, Yonelinas, & Ranganath, 2007; Squire, Stark, & Clark, 2004) and an absence of link between the STM/WM and the medial temporal lobe (for reviews, Baddeley, 2003; Collette, Hogge, Salmon, & Van Der Linden, 2006; D'Esposito, 2001).

Several authors have cast doubt on this 50 year old link by putting forward data that show that the medial temporal lobe could also be engaged in STM tasks (e.g., Ranganath & Blumenfeld, 2005; Ranganath & D'Esposito, 2005; and Nichols, Kao, Verfaellie, & Gabrieli, 2006). However, recent reviews have proposed a convincing explanation concerning this contradiction between 50 years of neuropsychology and some neuroimaging studies.

Jonides et al. (2008) made the observation that the neuroimaging studies in question have used tasks with longer retention intervals than those employed in neuropsychological studies, making STM tasks look more similar to LTM tasks. This crucial and basic idea that tasks must tap STM to uncover STM activity has also been put forward in a recent review by Squire and Wixted (2011). These authors claim that the medial temporal lobe does not play any kind of role in WM, and that these results are only an artifact. For these authors, it is only when WM capacity has been exceeded, and therefore when performance must also rely on LTM, that the medial temporal lobe becomes active. Squire and Wixted (2011) suggest that if methodological precautions are taken, then the results are straightforward as shown in recent studies (Jenison, Mauldin, Hopkins, & Squire, 2011; Jenison, Mauldin, & Squire, 2010; Shrager, Levy, Hopkins, & Squire, 2008).

Of course, an alternative possibility is that neurons subtending different cognitive functions are more closely interleaved than normally assumed (Smith & Kosslyn, 2007); this could for example be the case for the neural tissues devoted to WM and LTM in the medial temporal lobe. If the latter hypothesis turns out to be correct, more sophisticated techniques (such as high-resolution fMRI) would be needed to identify the two stages that we propose. But

it should also be noted that chunk-based theories, by assuming that pointers to LTM chunks are placed in WM, to some extent capture the close interaction between these two sets of structures.

However, building on Jonides et al.'s (2008) and Squire and Wixted's (2011) arguments and basing ourselves on the substantial neuropsychological and neuroimaging literature (see also Section 5.1.), we feel that the medial temporal lobe can be used as a signature for LTM. Doing so, we assume that it is possible to "reverse engineer" the mapping between brain structures and cognitive functions, arguing that reverse inference is the name of the game of much brain imaging. While techniques have been developed to facilitate such inferences (Schyns, Gosselin, & Smith, 2009), the logical and technical difficulties of doing so are considerable (Poldrack, 2006), and some authors have questioned the assumption that WM and LTM cognitive constructs can be mapped to specific regions in the brain (e.g., Postle, 2006). Nonetheless, building on Poldrack's (2006) argumentation, we will show that this issue is mitigated in our review.

4.2. Reverse inference and Bayes factor

Reverse inference can be translated into probabilistic terms (Poldrack, 2006; Sarter, Berntson, & Cacioppo, 1996) using Bayes' theorem. Doing so, it appears that the degree of belief in a reverse inference (in order to know if the activation of an area increases the probability of a process to have occurred, here LTM or WM) depends upon the selectivity of the neural response (viz., the ratio of process-specific activation to the overall likelihood of activation in that area across all tasks) and the prior belief in the engagement of cognitive process (here LTM or WM) given the task manipulation. It follows that if one wants to increase confidence into reverse inference, one needs to have greater selectivity of response in the brain region of interest, or/and a higher prior probability of the cognitive process in question.

Concerning the latter, a clear and independently supported theory seems a way to increase the prior probability of a cognitive process to be engaged. In our case, and taking experts undergoing WM-related tasks, LT-WMT (Ericsson & Kintsch, 1995) and TT (Gobet & Simon, 1996a) strongly predict the involvement of LTM processes. Predictions, it is to be noted, that were made before the data themselves were collected.

Concerning selectivity in reverse inference, if a region (e.g., the hippocampus) is activated relatively selectively by a specific process of interest (e.g., an LTM task), then one can infer with substantial confidence that the process is engaged given activation in the region. Poldrack (2006) sketches out a method to estimate selectivity using internet databases and the Bayes factor (i.e., the ratio of the posterior odds to the prior odds), in order to calculate if the activation in the area of interest increases the odds of engagement of the cognitive process. We followed the procedure using the same database (<http://www.brainmap.org/>; Laird, Lancaster, & Fox, 2005) as Poldrack (2006). Before taking a look at the results, it is important to mention that, according to Jeffreys (1961) and Kass and Raftery (1995), a Bayes factor between 1 and 3 offers weak evidence, between 3 and 10 offers substantial evidence, and from 10 offers strong evidence.

Taking the medial temporal lobe (the hippocampus and parahippocampal gyrus) as Location and WM as the Behavioral Domain as noted in the database, the Bayes factor is 0.4 (see for data, Appendix A, Table 1A); conversely, for explicit LTM as Behavioral Domain and for the same Location, the Bayes factor is 6 (see for data, Appendix A, Table 2A; the factor is above 10 for the hippocampus taken alone). This pattern is reversed, when one takes the prefrontal lobe (inferior, middle and superior frontal gyri) as Location. That is, the Bayes factor is above 10 (see for data,

Appendix A, Table 3A) when WM is the Behavioral Domain and 1.88 (see for data, Appendix A, Table 4A) when it is explicit LTM.

From this, two important things can be derived. First, since the degree of belief in a reverse inference depends partly upon the selectivity of the neural response and the prior belief in the engagement of a cognitive process (LTM or WM) given the task manipulation, we believe that in our case reverse inference can be used. Second, given the Bayes factors computed above, it appears that reverse inference gives us quantitative arguments to postulate that the medial temporal lobe is clearly linked with LTM while the medial temporal lobe is not linked with WM.

4.3. Physiological predictions

Putting all these remarks together, we believe that we can feel confident about the following two predictions. First, based on the Bayes factors computed above and the neuropsychological and neuroimaging literature (presented above and in Section 5.1.), we predict that medial temporal activation, which is supposed to represent episodic LTM activation, should be absent in novices in WM-related tasks while prefrontal and parietal activations should be found. Second, based on TT and LT-WMT, the Bayes factors, and the neuropsychological and neuroimaging literature (presented above and in the Section 5.1.), we predict that experts' brain activation will include a medial temporal activation in WM-related tasks and prefrontal and parietal activations. Therefore, the key element for contrasting both groups should be the medial temporal lobe.

Concerning semantic memory and the areas involved (mainly lateral and ventral areas of the temporal lobe, see for more precisions Section 5.1.), unfortunately the difference between experts and novices cannot be predicted in binary terms, absence vs. presence. In fact, in both groups, one can expect semantic knowledge to be activated during WM-related tasks, the areas depending on the type of information manipulated in the task (e.g., knowledge of words, knowledge of objects, knowledge of faces, and so on). The only prediction that could be made concerns the richness of the semantic knowledge involved, which should be more important for the experts in their domain of expertise; this could have a consequence in terms of brain activation clusters, but not in a clear-cut way as for the episodic memory.

5. Neuroimaging of experts

Chronologically, the first study using experts that exhibited results consistent with a cerebral functional reorganization implicating LTM areas was the one by *Pesenti et al. (2001)*. The authors contrasted an expert prodigy (R. Gamm, 6 year of practice) with a group of non-experts, asking all the participants to carry out simple multiplications (the scan base-line condition) and complex ones. While simple multiplications were supposed to be just retrieved, complex ones necessitated retrieval and multi-step calculations through WM. Due to R. Gamm's expertise, the step dynamics used in his multi-step calculations were highly structured and practiced (see *Pesenti, Seron, Samson, & Duroux, 1999; Pesenti et al., 2001*), taking the form of schemas or retrieval structures, LTM knowledge that allowed him to encode information in LTM as suggested by LT-WMT and TT.

The PET results (see Table 1 for results *in extenso*) showed that the non-expert group and the expert used some of the same brain areas: mainly prefrontal areas and parietal areas, with also occipital areas involved and the left inferior temporal gyrus. By contrast the areas activated only by the expert were the following: the right medial frontal gyrus, the upper part of the right anterior cingulate, the left paracentral lobule, the right middle temporal gyrus, and the right parahippocampal gyrus.

In the field of chess, *Saariluoma, Karlsson, Lyytinen, Teräs, and Geisler (2004)* asked six chess (37 years of practice) players to play blindfold chess and to execute blindfold chess tasks during a PET investigation. *Saariluoma et al. (2004)* did not employ a control group, because contrary to a normal chess game, a novice cannot play blindfold chess without being quickly overwhelmed. In order to play blindfold, experts must keep track of the location of each piece on the board, and therefore they have to continuously update their representation of the game via WM while thinking and planning the subsequent moves in WM. To do so, experts use also their knowledge of the game under the form of LTM representations: chunks (*Chase & Simon, 1973a*), templates (*Gobet & Simon, 1996a*) and retrieval structures (*Ericsson & Kintsch, 1995*).

The authors suggested that contrasting a memory condition (participants had to follow a chess game auditorally, memorizing the moves of a game and recalling them when prompted) and an attention condition (participants had to identify a specific piece when its name was pronounced) would provide information about mental imagery and storage in LTM (*Ericsson & Kintsch, 1995; Gobet & Simon, 1996a*). This is because none of these processes were required in the attention condition, while the memory condition necessitated storing the locations of the pieces and their moves (in addition to the identification process necessary in the attention condition).

PET results (see Table 1) showed mainly an activation pattern that concerned bilaterally the middle frontal gyrus, the angular gyrus bilaterally, and the inferior temporal cortex. There was also activation in the middle cingulate, and in the right middle temporal cortex.

The same kind of pattern of brain activation was also obtained with chess players in another study, but in a less constrained situation. *Campitelli, Gobet, Head, Buckley, and Parker (2007)* asked five chess experts to recognize previously-presented chess patterns during functional magnetic resonance imaging (fMRI) with a forced-choice procedure. Every sample stimulus was displayed for 6.5 s; this stimulus could belong to one of three experimental conditions: a "game" condition or a "random" condition (16 pieces displayed as a game position or randomly on a 4×8 matrix), a "scene" condition (16 geometrical figures displayed on a board the same size as the 4×8 matrix). After a 2-s delay, seven test stimuli of the same type appeared sequentially. For the first three conditions, the participant had to decide whether each one of the 7 test stimuli matched the sample stimulus. This is a standard WM task. However, experts can typically use simple patterns (chunks) and slotted schemas (templates) to encode and retrieve the incoming information using both WM and LTM (*Ericsson & Kintsch, 1995; Gobet & Simon, 1996a*).

Subtracting the scene condition (which was supposed to elicit WM but not chess LTM knowledge) from the game condition, and subtracting the random condition (which was supposed to elicit WM but not chess LTM knowledge) from the game condition had for result activation clusters in the following areas: posterior cingulate, both precuneii, right supramarginal gyrus, left lingual gyrus, temporal lobe areas including right fusiform gyrus, as well as inferior temporal gyrus and parahippocampal gyrus bilaterally.

Maguire, Valentine, Wilding, and Kapur (2003) used also fMRI to compare a group of 10 control participants to a group of 10 memory experts who had been using mnemonics techniques for 11 years. During the fMRI session, the experimenters asked the two groups to remember sequences of six items presented visually. Each item was presented alone for 4 s. It is during this period of presentation that the scans of interest were made. Three types of items were employed: three-digit numbers, faces and snowflakes. Just after the six items, participants saw three pairs of items from the last sequence and, had to indicate which item of each pair was presented first in the last sequence. This task typically elicits WM.

However, the experts, instead of simply storing the information in WM, used the method of loci.³ Therefore, they used a mental route as a LTM slotted schema associating the items to be remembered with the slots (the salient points or the familiar locations). The neuroimaging data showed, in both groups and during all stimuli (see Table 1 for results *in extenso*), an activation in frontal areas, in the left angular gyrus, in the left middle occipital gyrus, and in the caudate and cerebellum bilaterally. By contrast, three brain regions were active only for the mnemonist group for all stimuli types: the left medial superior parietal cortex, the bilateral retrosplenial cortex and the right posterior hippocampus.

We conclude this first part of our review by discussing expertise in mental abacus calculation. Just like Maguire et al. (2003) did with mnemonists, Chen, Hu et al. (2006) used fMRI to compare a group of (eight children) novices to a group of (nine children) experts. The experts had practiced operations and mental abacus calculations for 30–60 min a day, for about 5.5 years on average. Simple (single-digit numbers) and complex additions (four-digit numbers for the experts and two-digit numbers for the controls) were used. This task is demanding in terms of WM resources, especially with complex calculations. With experts, it is proposed that they use a mental representation of an abacus that is stored in LTM, which constitutes a retrieval structure (see Ericsson & Kintsch, 1995).

The fMRI results showed that in the case of novices, the activation pattern for both simple and complex calculations concerned the left inferior frontal gyrus, the anterior cingulate, the postcentral gyrus, and the inferior parietal lobule bilaterally. For novices and experts, activations were observed in the right superior frontal gyrus for simple calculations and in the left middle frontal gyrus and the left parahippocampal gyrus for complex calculations. In the case of experts and for simple calculations, activations were observed in the right lateral premotor cortex and bilaterally in the posterior temporal areas, and for complex calculations the activation occurred bilaterally in the lateral premotor cortex, the posterior superior parietal lobule, and in the parahippocampal gyrus.

5.1. Summarizing the cerebral activation in experts and providing evidence of a LTM involvement in WM-related tasks

The first part of Table 1 presents the neuroimaging studies we have just reviewed. When one examines the columns of Table 1 presenting the areas activated by novice controls and experts, it seems clear that prefrontal and parietal areas are involved to a great extent. This result was expected given the fact that the tasks used were WM-related and given the importance of these areas for WM (for prefrontal areas see for example, D'Esposito et al., 1995; Postle, Berger, & D'Esposito, 1999; Prabhakaran, Narayanan, Zhao, & Gabrieli, 2000; Shimamura, 1995; Ungerleider, Courtney, & Haxby, 1998; and for the importance of parietal areas see for example, Cowan, 2001; Cowan, 2005; Dehaene & Cohen, 1994; Postle & D'Esposito, 1999; Rodriguez et al., 1999; Todd & Marois, 2004).

By contrast and noticeably, the areas activated only in experts showed activations in medial temporal regions, such as the parahippocampus for Pesenti et al. (2001), Campitelli et al. (2007) and Chen, Hu et al. (2006) and in the right hippocampus for Maguire et al. (2003). As emphasized earlier, we believe that this is a marker of experts' strong reliance on episodic LTM during

WM-related tasks, which is also the point of view put forward by the authors of the studies reviewed.

Concerning the lateral and ventral temporal regions that we relate to semantic memory, experts showed more clusters of activation distributed across these areas. Activations were observed in the inferior temporal gyrus (right, Campitelli et al., 2007; or bilaterally, Chen, Hu et al., 2006; Saariluoma et al., 2004), in the middle temporal gyrus (bilaterally, Chen, Hu et al., 2006; or right Saariluoma et al., 2004), in the superior temporal gyri (bilaterally, Chen, Hu et al., 2006), and in the fusiform gyrus (right, Campitelli et al., 2007, or bilaterally, Chen, Hu et al., 2006).

In the following paragraphs, we provide evidence from the literature in favor of the involvement of the above-mentioned areas concerning LTM, starting with the medial temporal lobe and episodic memory. Activations in medial temporal structures (hippocampus and parahippocampus) are often observed during episodic encoding and retrieval (sometimes bilaterally), regardless of the verbal or non-verbal nature of the materials tested (for reviews, Eichenbaum et al., 2007; Squire et al., 2004). More specifically, the parahippocampal region in the right hemisphere seems to control the storage and maintenance of stimuli representations for long delays (Young, Otto, Fox, & Eichenbaum, 1997), and seems predominantly dedicated to the visuospatial aspects of these processes. It has also been proposed that this area plays an important role in a kind of expertise for local visual environment. Epstein and colleagues (Epstein & Kanwisher, 1998; Epstein, Stanley, Harris, & Kanwisher, 2000) have thus proposed the name of parahippocampal place area, but it is possible that in other kinds of expertise (for example in chess, Campitelli et al., 2007) the utilization of this area could slightly be transformed in order to be used in other kind of visual expertise (e.g., from local visual environment to chess).

The right posterior hippocampus has been shown to be very important for LTM and specifically for episodic memory. Several important models of memory are built around the hippocampus, making it play an essential role in the encoding and retrieval processes (the "standard" consolidation model, e.g., Squire & Alvarez, 1995; the multiple trace theory, e.g., Nadel & Moscovitch, 1997; the HIPER model, Lepage, Habib, & Tulving, 1998). The posterior hippocampus has been said to be more involved in the encoding of unfamiliar information (Gabrieli, Brewer, Desmond, & Glover, 1997) but conversely also in the retrieval process (Lepage et al., 1998). The right posterior hippocampus has also a clear involvement in spatial memory and navigation (Burgess, Maguire, & O'Keefe, 2002; O'Keefe & Nadel, 1978). In the case of Maguire et al.'s (2003) experts, this activation could be due to the use of the method of loci.

While these regions seem linked to episodic LTM, we believe that they are poorly linked to WM. And even if some researchers have exhibited a link between WM and the medial temporal lobe (e.g., Ranganath & Blumenfeld, 2005; Ranganath & D'Esposito, 2005; and Nichols et al., 2006), we believe that this evidence would have to be compared with the huge amount of neuropsychological and neuroimaging literature (for reviews, see Baddeley, 2003; Cave & Squire, 1992; Collette et al., 2006; D'Esposito, 2001; Gazzaniga et al., 2009; Müller & Knight, 2006; Squire & Wixted, 2011), where such a link has not been found. As seen in Section 4.2, one way of approaching this question is to compute Bayes factors in order to estimate the selectivity of the medial temporal lobe for each process (LTM vs. WM). It appears that there is no evidence of a link between WM and the temporal medial lobe while there is substantial evidence for a link between explicit LTM and the medial temporal lobe.

Concerning the lateral temporal cortex (superior, middle and inferior temporal gyri), it seems clearly involved in semantic memory, as shown since the late 1970s in clinical research in several pathologies (Alzheimer's disease, e.g., Martin & Fedio, 1983; Pick's

³ This famous mnemonic technique has been known for at least 2500 years. The ancient Greeks used it to remember important words during speeches. Before a speech, an orator would visualize a familiar route – which can be considered as a schema – with salient points along it or familiar locations (loci means "locations" in Latin) and use them to mentally store or associate important words. Later, during the speech, the orator could retrieve these words just by mentally walking down the route again.

disease, e.g., Warrington, 1975; semantic dementia, e.g., Hodges, Patterson, Oxbury, & Funnell, 1992; and strokes, e.g., Cloutman et al., 2009). More recently the link between temporal atrophy (especially in the anterior part) and semantic memory deficits has been confirmed by several authors (Galton et al., 2001; Garrard & Hodges, 2000; Gorno-Tempini et al., 2004; Mummery et al., 2000). Neuroimaging studies have also confirmed the involvement of the lateral temporal cortex in semantic memory, especially in the left hemisphere (e.g., Binder et al., 1997, 1999; Chee, O'Craven, Bergida, Rosen, & Savoy, 1999; Démonet et al., 1992; Mummery, Patterson, Hodges, & Price, 1998; Vandenberghe, Price, Wise, Josephs, & Frackowiak, 1996). Even if the majority of the studies favor the hypothesis that semantic information is stored or processed mainly in the left side of the brain (for a review, see Martin, 2006), the exact organization of semantic memory does not seem to be perfectly known, within each hemisphere (for an overview, see Caramazza & Mahon, 2006) or even when comparing one hemisphere to the other (e.g., Damasio, Tranel, Grabowski, Adolphs, & Damasio, 2004).

As far as the ventral areas of the temporal lobes (the inferior temporal and the fusiform gyri in the studies we reviewed) are concerned, cognitive neuroimaging studies in humans and single-cell recording in non-human primates seem to indicate that they are involved in the storage of familiar patterns in LTM (e.g., Damasio, Grabowski, Tranel, Hichwa, & Damasio, 1996; Desimone, Albright, Gross, & Bruce, 1984; Gross, 1992; Logothetis, Pauls, & Poggio, 1995; Stark & Squire, 2000; Tanaka, 1993). In the right side of the brain ventral areas, the fusiform gyrus and the inferior temporal gyrus are known to be involved in object recognition, and more generally in the encoding and retrieval of the figurative properties of visual representations, even when the objects are just imagined (e.g., Mellet et al., 1996).

As for the fusiform gyrus in particular – most often the right mid-fusiform, but sometimes bilaterally – a subset of it has been named the Fusiform Face Area (FFA), because (a) its activity increases when faces are seen or generated by mental imagery (Kanwisher, Tong, & Nakayama, 1998; McCarthy, Puce, Gore, & Allison, 1997; O'Craven & Kanwisher, 2000); and (b) because its response is greater for faces than for objects, animals or scenes, or other parts of the human body (Aguirre, Singh, & D'Esposito, 1999; Haxby, Hoffman, & Gobbini, 2000; Ishai, Ungerleider, Martin, Schouten, & Haxby, 1999; Kanwisher, McDermott, & Chun, 1997; Kanwisher et al., 1998; Yovel & Kanwisher, 2004). However, the FFA may not be specifically a face area. Gauthier, Tarr, and Anderson (1999) suggested that the FFA may be involved in visual processing for items individuals have expertise in: faces, of course, but also cars and birds for car and bird experts (e.g., Gauthier, Skudlarski, Gore, & Anderson, 2000). Therefore, taking these results into account, it is possible that in the case of Campitelli et al.'s (2007) chess experts and Chen, Hu et al.'s (2006) child abacus experts, the activation of the right fusiform (bilaterally for Chen, Hu et al., 2006) could be due to a form of visual processing and mental imagery for items for which these individuals have expertise.

Taken together, these results seem to indicate that the areas used by experts in WM-related tasks tend to be, in addition to the WM areas used by novices, also LTM areas. This is clearly in line with LT-WMT (Ericsson & Kintsch, 1995) and TT (Gobet & Simon, 1996a), which, in the case of experts executing WM-related tasks, both hypothesize the use of LTM via knowledge structures.

5.2. Neuroimaging of experts: strong hypothesis vs. weak hypothesis

If the first part of Table 1 presents data that is directly compatible with LT-WMT (Ericsson & Kintsch, 1995) and TT (Gobet & Simon, 1996a), the three studies presented in the second part of Table 1 (Chen, Wu et al., 2006; Hanakawa et al., 2003; Tanaka,

Michimata, Kaminaga, Honda, & Sadato, 2002) do not exactly fit in. In fact, instead of having data compatible with a functional reorganization via an involvement of episodic LTM areas (medial temporal lobe) or semantic LTM areas (ventral and lateral temporal areas), these studies are only compatible with a functional reorganization non-specific to LTM. By “functional reorganization non-specific to LTM,” we mean that when an individual passes from the status of novice to that of expert, there is indeed a change concerning the pattern of areas involved with the activation of new brain areas or the deactivation of old brain areas (viz., functional reorganization) but this change does not concern the involvement of more LTM areas (LTM non-specific).

Ericsson (2003, p. 235, text in brackets added) has clearly stated that both types of results (a functional reorganization with involvement of LTM areas and a non-specific functional reorganization) are compatible with LT-WMT, and we agree with him: “Most important, the differential brain activation (between experts and a control group) is consistent with cognitive processes predicted by long-term working memory accounts. For example, exceptional mental calculators rely on storage in long-term memory (here Ericsson refers to Pesenti et al. (2001)) and expert mental abacus calculators encode numbers in a manner qualitatively different from controls (here Ericsson refers to Tanaka et al.'s (2002) study, which we present below).

Even if we agree with Ericsson about the compatibility of both results, we feel that the latter result (a non-specific functional reorganization) is qualitatively more distant from LT-WMT and TT compared with a brain functional reorganization with involvement of LTM areas, since expertise is supposed to be reached through the building of LTM structures (chunks and knowledge structures). Therefore we have decided to name the non-specific functional reorganization the “weak physiological hypothesis” (“weak hypothesis” from here on) and the brain functional reorganization with involvement of LTM areas “the strong physiological hypothesis” (“strong hypothesis” from here on). We next review and explain the studies showing results compatible with the weak hypothesis.

5.3. Three abacus studies in favor of the weak hypothesis

To our knowledge, Tanaka et al. (2002) was the first study examining abacus experts with fMRI neuroimaging. The authors compared ten abacus experts that had from 8 to 16 years of training to thirteen non-expert controls. A delayed match-to-sample task was utilized during fMRI. A target sequence was presented during 3 s; for each participant, the length of the digit sequence was two digits shorter than their digit span score (measured previously). After a 15-s delay participants had to decide whether a test sequence was the exact same sequence as in the sample. This task is typically supposed to tap WM. However, as in Chen, Hu et al. (2006), experts were also supposed to use their LTM through their mental abacus, which constitutes a retrieval structure that helps them encode information in LTM (see Ericsson & Kintsch, 1995).

Regarding the fMRI data of the control group, the areas engaged were located mostly in the left hemisphere (Broca's area, insula, medial and precentral gyrus), plus bilaterally the parietal lobules and the cerebellum. For the experts, the areas engaged were symmetrical, mainly in the superior frontal sulcus (bilaterally) and in the superior and inferior parietal lobules (bilaterally).

Similar results were observed by Hanakawa et al. (2003), which have also tested abacus experts in an fMRI experiment, using WM-related tasks. They compared six experts that had more than 17 years of almost daily training to eight controls. All participants performed a numeral mental-operation task as in Chen, Wu et al. (2006). Sixteen numbers were presented, one every 2 s. Participants had to sequentially add the numbers and keep in mind the final result, in order to perform a recognition test.

For the controls, except for a bilateral activation in the superior precentral sulcus and the cerebellum, all the other activations were in the left hemisphere: in the prefrontal cortex, in Broca's area, in the medial frontal gyrus, in the inferior parietal lobule, in the precuneus, in the intraparietal sulcus, in the fusiform gyrus. For the experts, except for a left activity in the fusiform gyrus, results in the numeral task showed a bilateral activity: in the inferior frontal gyrus, in the superior precentral sulcus, in the precuneus, in the intraparietal sulcus, and in the cerebellum.

The final study in this section is Chen, Wu et al.'s (2006) experiment. Like in the two studies we have just reviewed, Chen, Wu et al. (2006) carried out an fMRI investigation using abacus experts. The authors asked six experts and six controls to execute two types of calculations: one-digit additions or two-digit additions. Covert reading (of the numbers to add) was used as the control task. The fMRI results were obtained subtracting the covert reading from the calculation condition.

For the non-experts, one-digit and two-digit calculations engaged mainly the same left hemisphere areas with an exclusive activation of prefrontal and parietal areas. The activations were in the left Brodmann areas (BA) (Broca's area) 44 and 45⁴ and in the left BA 39. In addition the left BA 6 and BA 7 and 40 were also activated (on both sides of the brain during two-digit calculations). Finally, the BA 24 and 32 were also activated.

In the case of the experts, the pattern of activation was symmetrical, exactly like in Tanaka et al.'s (2002) and Hanakawa et al.'s (2003) studies. In the two-digit condition, there was a bilateral activation of the precentral gyrus and in parietal areas (BA 7, 40 and 19). The right postcentral gyrus (BA 2) was also activated. In the two-digit condition, the same pattern was found, with an activation in the postcentral gyrus (BA 3 and 4).

5.4. Explaining the weak hypothesis in abacus studies

The three abacus studies we have just reviewed suggest that the cerebral areas used by abacus experts – during additions in Chen, Wu et al. (2006) and Hanakawa et al. (2003), or during a digit span in Tanaka et al. (2002) – are different from those used by control novices in these same tasks (see also Table 1). Even if frontal and parietal areas are involved in both cases, the pattern is symmetrical for experts.

All the authors have suggested that in the case of experts, the pattern is compatible with visuospatial/visuomotor imagery processing (e.g., Courtney, Petit, Maisog, Ungerleider, & Haxby, 1998; Rowe, Toni, Josephs, Frackowiak, & Passingham, 2000), whereas for controls, the pattern is more compatible with verbal processing of information (e.g., Burbaud et al., 1999; Dehaene et al., 1996; Pesenti et al., 2001; Zago et al., 2001).

This result is clearly congruent with a functional brain reorganization that would occur during expertise acquisition – the weak physiological hypothesis derived from LT-WMT and TT – but does not seem to be compatible with a functional brain reorganization involving LTM areas, the strong hypothesis.

The question that stems from these results is as follows: why are all studies using abacus experts only compatible with the weak hypothesis (with the exception of Chen, Hu et al. (2006), while the studies involving other tasks (mental multiplication, blindfold chess, chess positions recognition, digits, faces and snowflakes

serial order recognition) are consistent with the strong hypothesis? We think that an important part of the answer is to be found in the dissociation between the areas involved in general semantic LTM and the areas involved in semantic LTM concerned specifically with the representation of numbers.

As seen previously, neuroimaging studies have shown that general semantic LTM tasks tend to activate mostly a network of areas where the left temporal cortex plays a crucial role (e.g., Martin, Haxby, Lalonde, Wiggs, & Ungerleider, 1995; Martin, Wiggs, Underleider, & Haxby, 1996; Mummery et al., 1998; Tyrrell, Warrington, Frackowiak, & Rossor, 1990). By contrast, the sense of number seems to activate parietal areas, near the intra-parietal sulcus and neighbor areas (like the inferior parietal lobe). Some studies pinpoint the sense of number only in the left parietal cortex (Chochon, Cohen, Van de Moortele, & Dehaene, 1999; Cowell, Egan, Code, Harasty, & Watson, 2000) while other studies also localize it in the right parietal cortex (Chochon et al., 1999; Dehaene et al., 1996; Kazui, Kitagaki, & Mori, 2000). Studies using single neuropsychological cases have yielded the same kind of results showing a double dissociation between general semantic knowledge and numerical semantic knowledge. For example, in cases of semantic dementia (Warrington, 1975) where there is an atrophy of the temporal lobes (e.g., Garrard & Hodges, 2000; Hodges et al., 1992; Mummery et al., 2000), general semantic knowledge is gradually lost while numerical capacities are spared (e.g., Cappelletti, Kopelman, Morton, & Butterworth, 2005; Crutch & Warrington, 2002). Inversely, general semantic knowledge can be spared while numerical capacities are lost in Gertsman syndrome (Gerstmann, 1940) or in posterior cortical atrophy (e.g., Delazer, Karner, Zamarian, Donnemiller, & Benke, 2006; Tang-Wai et al., 2004), which both concern mainly parietal areas.

Therefore, it is possible that, in the case of abacus experts, there is indeed a functional brain reorganization involving LTM areas, like in the other studies we have reviewed but, because of the peculiar cerebral localization of numerical semantic knowledge in the parietal lobes, no physiological reorganization is apparent. In fact, the parietal areas (mainly the posterior areas) are important for both WM (e.g., Postle & D'Esposito, 1999; Rodriguez et al., 1999; Todd & Marois, 2004; for the importance of parietal areas in WM, see also Cowan, 2001; Cowan, 2005) and for numerical semantic knowledge (e.g., Chochon et al., 1999; Cowell et al., 2000; Dehaene et al., 1996; Kazui et al., 2000). Therefore, even though with expertise acquisition, in the case of abacus experts, there is a gradual activation shift from WM areas (prefrontal and parietal) to WM areas (prefrontal and parietal) and LTM areas (parietal in the case of mental abacus calculations), it is understandable that this shift goes undetected, since parietal areas are crucially involved in both cases (WM and LTM).

The results of the three above-mentioned studies (Chen, Wu et al., 2006; Hanakawa et al., 2003; Tanaka et al., 2002) clearly confirm the involvement of parietal areas, since it appears that the only areas that are common to the experts in all three studies are the intra-parietal sulcus and neighboring areas. More precisely, these areas are the intra-parietal sulcus (bilaterally) in Hanakawa et al. (2003), the two parietal lobes (bilaterally, superior and inferior) that surround the intra-parietal sulcus in Tanaka et al. (2002), and BA 40, 7 and 19 bilaterally in Chen, Wu et al. (2006). Moreover, in the three studies when the experts - controls contrast was statistically tested, the only areas (plus the superior frontal sulcus in Tanaka et al., 2002) that showed a significant activation difference in favor of the experts were in the parietal region, and more precisely the superior parietal lobule in Tanaka et al. (2002) and Hanakawa et al. (2003), and the postcentral gyrus in Chen, Wu et al. (2006).

Given these findings, we suggest that the results observed by Tanaka et al. (2002), Hanakawa et al. (2003) and Chen, Wu et al.

⁴ In their study, Chen, Wu et al. (2006) did not always provide the stereotaxic coordinates and the names of the gyri that were activated; instead they always indicated the Brodmann areas involved. In our review, we have always indicated the gyri and/or sulci, but since the overlap between Brodmann areas and gyri is not perfect and since the stereotaxic coordinates have not been provided, when it was not possible to do otherwise we have just indicated the Brodmann areas concerning Chen, Wu et al.'s (2006) study.

(2006) can also be consistent with a brain functional reorganization involving LTM areas, although these are numerical semantic LTM areas.

6. Neuroimaging of trained novices

The results from the fMRI and PET studies that we have reviewed above in different tasks and in different domains clearly suggest that experts and novices do not utilize the same brain areas in tasks related to WM. We have argued that LTM areas are strongly involved in the case of experts (for a summary of the results, see Table 1) and not in the case of novices. We consider this finding as evidence in favor of a functional cerebral reorganization involving LTM areas related to expertise acquisition. However, we acknowledge that the evidence is only *indirect* because two different populations are being compared. In principle, a better approach would be to follow the evolution of expertise acquisition from the start in a group of novices, and look for brain functional reorganization. Some studies have followed this approach. But, surprisingly, when novices have been trained to gain expertise in tasks related to WM, no cerebral functional reorganization has been observed, the results mostly showing a decrease in activation.

We will next review these studies on novice training, using two dimensions to group them. The first dimension concerns the experimental paradigm employed and the second concerns the duration of practice. With regard to the experimental paradigm, the great majority of the authors (eight studies out of eleven) have used what we have called the Scanning–Training–Scanning (S–T–S) paradigm, where the brain activity is measured before training and after training, or additionally in between, but the main idea is to compare the situation before training to the situation after training. By contrast, in three studies (Andreasen, O'Leary, Arndt, et al., 1995; Andreasen, O'Leary, Cizadlo, et al., 1995; Moore, Cohen, & Ranganath, 2006), the authors have used what we have called the Training–Scanning (T–S) paradigm, where participants are first trained concerning a task and a type of material and scanned only after. We first present the studies using the S–T–S paradigm; then, we describe the T–S paradigm, presenting the three studies that are concerned.

With regard to the practice duration dimension, short training regimens (in the order of hours) were employed in six studies, while medium-length training regimens⁵ (in the order of weeks) were employed in five studies. We next review these studies, beginning with short-training regimens. It is to note that we have taken into consideration for the short-length training only experiments that encompassed a practice of at least 1 h, which seemed to be an adequate length in order to start speaking of training.

6.1. Scanning–Training–Scanning paradigm experiments with short-length training

Chronologically, the first study using short training was Garavan, Kelley, Rosen, Rao, and Stein's (2000). They used a visuospatial delayed-match-to-sample task that was used originally by Jonides et al. (1993). During the presentation phase, participants saw three dots briefly. After a 3-s delay, three dots were again displayed with a probe circle surrounding one of those dots. Participants had to decide, by pressing a yes or no button, if the surrounded dot was at the same location as in the presentation phase. Participants had to complete 880 trials (over 8 h). Response times improved whereas accuracy did not. The decrease in response time corresponded to

a decrease in activation in the prefrontal and parietal areas that are involved in WM (Cabeza & Nyberg, 2000; Carpenter, Just, & Reichle, 2000) and in the occipital cortex; no functional reorganization was detected (see Table 2 for more details).

Landau and colleagues (Landau, Garavan, Schumacher, & D'Esposito, 2007; Landau, Schumacher, Garavan, Druzgal, & D'Esposito, 2004) also showed an activation decrease in delayed-match-to-sample tasks, using event-related fMRI, which allowed them to differentiate the encoding process from the retrieval and isolate the delay period. In the first study (Landau et al., 2004), a stimulus set composed of four intact or scrambled faces was shown to the participants, who were asked to remember all the intact faces after an 8-s delay period; there could be one, two, three or four intact faces (the load factor). During the scanning session, the first 18 min were compared to the last 18 min, for a total of 48 min of scanning (plus 30 min of unscanned practice). The results showed a decrease in activation in the frontal, temporal, parietal, and occipital areas related to practice, and no functional reorganization. Concerning the load factor and the periods of image acquisition (encoding, delay, and retrieval), a decrease in activation and effects of memory load both occurred primarily during encoding. The authors concluded that practice improved encoding efficiency, especially at higher memory loads. It is also important to note that the decreases in activation did not correspond to improvements in the behavioral task.

More recently, Landau et al. (2007) have confirmed the practice-related decreasing pattern but with a visual and a spatial WM task. Participants practiced approximately 60 min. Like in their previous study, the authors used an event-related fMRI investigation, but unlike the previous studies, they used irregular polygons, designed to be difficult to encode verbally. For spatial runs, participants were asked to remember the spatial position of polygons whereas for object runs, participants were asked remember the shape. Behavioral results showed a linear decrease of the reaction times during scanning. The accuracy did not change across the runs. With respect to the neuroimaging data, practice-effects were only observed during WM encoding and retrieval, but not during maintenance. The decrease of activation concerned mostly WM frontal and parietal areas (see Table 2 for more details).

Like Landau et al. (2007), Sayala, Sala, and Courtney (2006) have investigated the effect of practice on WM using spatial and object instructions, but they utilized faces as stimuli. In the sample phase a sequence of three stimuli was presented. After a memory delay of 3, 4.5 or 9 s, the test phase began; a face was displayed during 3 s. In object trials, participants had to decide if the face was part of the sample phase; in spatial trials, participants had to decide if a face in the sample had been displayed in the same location. The behavioral results showed no significant changes in terms of reaction times or errors across the scanning. Moreover, no correlations between fMRI activity and behavioral measures were observed.

The fMRI results were analyzed separately for the sample, the delay and the test phases. Across the three phases, and collapsing data across object and spatial trials, a decrease of brain activation was observed. The practice-related effect was more important during the sample and delay phases. The main network of regions common to the sample phase and the delay phase is congruent with WM activities: it mainly comprised frontal and parietal areas, plus the cingulate, and the insula. During the test phase, WM areas also were elicited: the left inferior frontal gyrus showed a practice-related decrease in activity (for more details see Table 2).

A decrease in activation in WM areas was also observed with a more complex task, which is mainly supposed to tap into the planning function. With a modified, computerized version of the tower of London, Beauchamp, Dagher, Aston, and Doyon (2003) wanted to observe the effect of repeated practice, that is, solving problems for 2 h. The training resulted in an improvement in behavioral

⁵ We consider a delay of weeks as a medium-length training, because the experts that are usually compared to novices in standard experiments (e.g., in the studies we reviewed in the previous section) go through much longer training periods, in the order of years.

performances: a decrease in the time required for planning and execution, as well as in the total time to solve the problems. Like in the previous studies, a decrease in frontal regions was detected. The authors observed a decrease mainly in the orbitofrontal and frontopolar regions (for more details see Table 2), which are particularly important for guessing, decision-making and use of feedback (Bechara, Damasio, & Damasio, 2000; Elliott et al., 1997; O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001; Rogers et al., 1999). The results also showed a positive correlation between activity in these areas and the decrease in time taken to solve the problems.

Jansma, Ramsey, Slagter, and Kahn (2001) used a verbal Sternberg task. Participants had to remember a target set composed of five consonants; then a consonant was presented and the participants had to decide whether it belonged to the target set; for each target set, 10 consonants were presented. Participants were asked to practice the task for 45 min and were then scanned for approximately 40 min while performing the same task. Behavioral data revealed that the answers became faster and more accurate. The fMRI data showed a decrease in activity during the scanning, mainly in areas related to WM: the bilateral dorsolateral prefrontal cortex, the right superior frontal cortex, the right frontopolar area and the supplementary motor area.

6.2. Scanning–Training–Scanning paradigm experiments with medium-length training

We review three studies that used medium-length sessions: 4–5 weeks. Using a spatial *n*-back task, Hempel et al. (2004) observed first an increase in activation in fMRI signals and then a decrease in WM areas. Participants were scanned over three periods: before training, after 2 weeks of training and after 4 weeks of training. A load factor was used by modifying the *n*-value of the *n*-back task (0-back, 1-back, and 2-back). Participants trained independently at home, twice a day. The fMRI results mainly showed an activation involving the right inferior and medial frontal gyri, the right intraparietal sulcus and the right superior parietal lobe. The load factor was associated with activation increases in these regions. The authors suggested that the increase–decrease pattern of activations followed an inverse *U*-shaped quadratic function.

Olesen, Westerberg, and Klingberg (2004) also observed an increase in brain activation in WM areas. They used a visuospatial WM task, training their participants for 5 weeks during 35–45 min of training per day. In their first experiment, brain activity was measured with fMRI twice, before and after training. In a second experiment, subjects were scanned once a week during a 5-week training period. Both experiments showed an increase in activity in the middle frontal gyrus and in the superior and inferior parietal cortices, correlated with an increase in WM performance due to training.

The important point concerning training was that *the difficulty of the training tasks was automatically adjusted* to performance by changing the number of stimuli to be remembered. This may explain, in different ways, the peculiar pattern observed, that is, brain increase activation. In fact, as pointed out by Kelly and Garavan (2005, p. 1097), “a number of studies have demonstrated how increased task difficulty or load is associated with increased BOLD activation (Rypma & D'Esposito, 1999; see also e.g., Druzgal & D'Esposito, 2003; Rypma, Berger, & D'Esposito, 2002).” The same authors also proposed a second explanation. As discussed above, Hempel et al. (2004) observed first an increase (0–2 weeks) and later a decrease (2–4 weeks) in activation in an *n*-back task. It is possible that Olesen et al.'s (2004) result could also be described by a quadratic function, since the final scanning session activation appeared to be decreasing (see Fig. 3d in Olesen et al., 2004).

Dahlin, Stigsdotter-Neely, Larsson, Bäckman, and Nyberg (2008) used the same training procedure, adjusting the difficulty of the training tasks to participants' performance. However, they adapted

it to an updating task. One group of 15 participants was trained during 5 weeks and compared with a group of 7 participants that did not receive any kind of training. Letters were employed in the updating task during fMRI acquisition. It is important to note that unlike all the previous studies, the authors did not only analyze a simple effect (pre-training vs. post-training), but also an interaction that incorporated the control group vs. training group contrast, comparing the training group pre-training vs. post-training difference with the control group pre-training vs. post-training difference.

Behavioral results showed that the training group exhibited considerably larger gains in both tasks than the control group. Concerning the fMRI data, an increase of activation was observed in the temporal lobe, in the striatum, and in the occipital lobe. Conversely a decrease was observed in the frontal and parietal lobe. These results could be compatible with a functional reorganization as defined earlier (Kelly & Garavan, 2005; Poldrack, 2000). However no behavioral indicators was available in the papers to confirm that with the increase of activation of new areas there also was a shift in the cognitive process underlying task performances. Finally it is interesting to note that among the regions that increased, a large cluster in the temporal lobe was found, which could be compatible with LT-WMT and TT predictions.

6.3. Training-Scanning paradigm experiments with medium-length training

We finish this review with three studies (Andreasen, O'Leary, Arndt, et al., 1995; Andreasen, O'Leary, Cizadlo, et al., 1995; Moore et al., 2006) that also used medium-length training, but in a different way compared with the studies we have just reviewed. In the eight previous studies, the authors have used the S–T–S paradigm comparing brain activation before training and after training. In contrast, in the three following studies, participants were first trained on a task and a type of material, and only after scanned (that is, these studies used the T–S paradigm). In order to have a contrast, the participants' brain activation was measured in two ways: (a) while the participants executed the task for which they had been trained for, and with the same material used during the training, the practiced condition, and (b) while the participants executed the task for which they had been trained for, but with a novel kind of material for which they had no expertise, the (pseudo)novice condition.

In a PET study, Andreasen and colleagues used the T–S paradigm to investigate the effect of practice with a narratives recall task (Andreasen, O'Leary, Arndt, et al., 1995) and with a word span memory task (Andreasen, O'Leary, Cizadlo, et al., 1995; in both cases, a practiced condition was contrasted with a novice condition). In Andreasen, O'Leary, Arndt, et al. (1995), two stories (Story A and Story B) were taken from the Wechsler Memory Scale-Revised (Wechsler, 1987). Participants heard Story A and recalled it continuously until they knew it perfectly; this was done 1 week before the scanning and one day before the scanning. Finally, just before the scanning, participants heard Story A again, then they had to recall Story A during the scanning process. Conversely, Story B was heard only once, just before the scanning, and then the participants had to recall it during the scanning. Behavioral data showed that participants recalled 99% of Story A and 56% of Story B. The structure of Andreasen, O'Leary, Cizadlo, et al.'s (1995) study was the same, except that they used two lists words (List A and List B) from the Rey Auditory Verbal Learning task (Rey, 1964). Participants scored a mean of 14.2 hits out of 15 words for List A and 6.6 hits for the 15 words of List B.

Concerning the PET signals, results from the two studies showed that the practiced condition produced smaller and fewer activations. When subtracting the practiced condition to the unpracticed condition, Andreasen, O'Leary, Arndt, et al. (1995) ob-

served activation in the left frontal operculum, in the cingulate gyrus, and in the cerebellum. In Andreasen, O'Leary, Cizadlo, et al.'s (1995) study, the same subtraction showed an activation of superior and inferior frontal areas, in the parietal region, and in the cerebellum. Results, especially from the latter study, are comparable with the main pattern observed in the studies using the S–T–S paradigm: a brain decrease activation pattern mainly in WM prefrontal and parietal areas, with no functional reorganization.

Moore et al. (2006) have also employed the S–T paradigm, in a match-to-sample task with complex visual objects (polygons). Participants were first trained during 10.5 h with two categories of objects. Half the participants were trained with one category and the other half with the second category. Once the training phase finished (an average of 10 days), participants were scanned during a match-to-sample task. In each trial, a cue (a complex object) was displayed followed by a delay period that ranged from 6.75 to 12.75 s. Finally a probe object was displayed for 1.25 s. Participants had to indicate if the cue and the probe were the same. Two types of objects were displayed: those that belonged to the category for which the participants had become experts (the practiced condition) and those that belonged to the category for which the participants were not trained (the novice condition).

Results showed that participants had a better accuracy in the match-to-sample task in the practiced condition, that is, for the objects they were trained with. The fMRI results showed mostly an increase of activation in WM frontal and parietal areas and in the middle occipital gyrus (for more detail, see Table 2). These activations areas are comparable with the main activation pattern observed in the studies using an S–T–S paradigm; however, in this last study, no activation decrease was observed, but instead an increase of activation.

6.4. Summary

In this section, we have reviewed twelve PET and fMRI studies in which novices were trained in order to observe brain activation changes during WM-related tasks (see Table 2). Among the studies using the S–T–S paradigm, six showed a decrease in activation in WM areas, one showed an initial increase followed, after 2 weeks, by a decrease in activation (Hempel et al., 2004), one showed only an increase of activation (Olesen et al., 2004), and one was compatible with a functional reorganization (Dahlin et al., 2008). Concerning the S–T studies, the two studies of Andreasen, O'Leary, Arndt, et al. (1995) and Andreasen, O'Leary, Cizadlo, et al. (1995) showed a decrease of activation, and Moore et al. (2006) observed an increase mainly in WM areas. We stress the fact that these three S–T studies cannot be directly compared to the studies using the S–T–S paradigm.

From the target studies we have reviewed in this section, it first appears that, with the exception of Dahlin et al.'s (2008) study, no data consistent with functional reorganization has been detected with WM-related tasks. Second, the main pattern observed is a practice-related decrease of brain activation, that is, eight among twelve. And third, the main areas engaged and showing a decrease of activation are – as expected given the tasks employed, that is WM – prefrontal and parietal areas. Table 2 presents all the studies we have reviewed in this section with all the areas showing practice-related changes.

7. Bridging the two patterns of neuroimaging

7.1. A two-stage view of expertise acquisition

Based on the description of the state of the art we have provided above, there seems to be a clear discrepancy between (a) the data

collected with novices who are trained to become experts, which mainly show a decrease of cerebral activity in WM areas, and (b) the data collected with experts who are compared to novices, which are compatible with a cerebral functional reorganization from WM areas to WM and LTM areas. We think that an important factor explaining this discrepancy could be the duration of practice and its content.

With regard to duration, several training periods in the former studies (showing mainly a decrease in cerebral activity) are probably too short – experiments using short-training regimens, in order of hours – to allow novices to build crucial cognitive structures for expertise, that is, chunks but especially knowledge structures (Ericsson and Kintsch's retrieval structures and Gobet and Simon's templates). Among the six studies using a medium-length training, which should be more favorable to functional reorganization, only one exhibited this pattern.

With regard to the content of the training sessions in the former studies (showing mainly a decrease in cerebral activity), and especially in the case of longer practice durations – experiments using medium-length training regimens, in order of weeks – the training does not seem enough oriented towards the acquisitions knowledge structures. Concerning this last point, it is to be noted that specific training programs (e.g., Ericsson & Harris, 1990; Gobet & Jackson, 2002; Saariluoma & Laine, 2001) or suggestions of mnemonics to participants (Chase & Ericsson, 1981, see Fig. 5.13) can speed up the process of acquiring chunks and especially knowledge structures. These last structures are crucial for expertise, because they allow individuals to expand their WM, enabling them to use part of their LTM as WM, which is theoretically supposed to cause a cerebral functional reorganization.

Even if we argue that the duration of practice and its content are crucial factors for explaining the difference between the data collected with novices and the data collected with experts, we do not deny the possibility that part of the difference between the two patterns could be due to the type of task employed. In fact when considering the tasks used in the neuroimaging studies of trained novices, they appeared to be often abstract, unlike the tasks used in the neuroimaging studies of experts. Some are constructed so that it is very difficult to use any semantic knowledge structures that may help out with doing the task. However, this is not the case for all the studies. In fact, Landau et al. (2004) and Sayala et al. (2006) used faces; Jansma et al. (2001) used letters; and Andreasen, O'Leary, Arndt, et al. (1995) and Andreasen, O'Leary, Cizadlo, et al. (1995) used words and sentences. That is why our claim is that the length and the content of the training period are the crucial factors explaining the discrepancy.

Our hypothesis is that the two brain-imaging patterns of results observed in the literature are not contradictory but are two stages of the same process: the acquisition of expertise. We propose that, in terms of brain activation, expertise acquisition in WM-related tasks is a two-stage process that starts with a decrease in activity and ends with a functional brain reorganization.⁶

As far as the first stage – brain decrease activation – is concerned, the only three studies that are not congruent with this view, among the twelve reviewed, are Hempel et al.'s (2004), Olesen et al.'s (2004), and Moore et al.'s (2006). As argued above, we considered the latter one with caution because of the paradigm used (T–S), which makes delicate the comparison with the majority of the studies that involve the S–T–S paradigm.

⁶ Kelly and Garavan give the following description of functional reorganization (2005, p. 1089, italics added): "Three main patterns of practice-related activation change can be distinguished. Practice may result in an *increase* or a *decrease* in activation in the brain areas involved in task performance, or it may produce a *functional reorganization* of brain activity which is a combined pattern of activation increases and decreases across a number of brain areas."

Concerning [Hempel et al.'s \(2004\)](#), [Olesen et al.'s \(2004\)](#) studies, they tend to show that the first stage of expertise could be an increase in brain activation, which is compatible with a three-stage process view of expertise: increase, decrease, and then brain functional reorganization. Nevertheless, we do not propose a three-stage view because, as we have already noted, [Olesen et al.'s \(2004\)](#) methodology is peculiar and could be misleading. As for [Hempel et al.'s \(2004\)](#) results, the increase in activity seems genuine but we think that more results of this kind are necessary before we should take into consideration an increase in activity with WM-related tasks.

In 2004, [Jonides](#) already hypothesized that the two patterns of activity described above could be two different consequences of practice. In his view, the decrease in activity could correspond to greater skill at applying the initial strategy and the functional reorganization could correspond to the development of new and more efficient strategies. However, [Jonides \(2004\)](#) did not specify when these patterns should appear during expertise acquisition, nor whether they should follow a sequential order.

The interpretation of the review we have provided is that the patterns occur in a sequence, with functional brain reorganization occurring after activation decrease. However, we do not think that this means that novices in the early stages of expertise acquisition do not develop new and more efficient strategies, as suggested by [Jonides](#). And, indeed, in the only study ([Olesen et al., 2004](#), experiment 2) that reported the content of participants' interviews about their strategy, two out of eight participants said that they had changed strategies. What we suggest, and what the data seem to indicate, is that a simple change of strategy is insufficient to induce functional brain reorganization in a context of practice: it may be a necessary but not a sufficient condition.

From the data we reviewed in the "Neuroimaging of experts" section, it seems that functional reorganization corresponds to a profound cognitive modification that occurs only at a late phase of expertise acquisition when important cognitive structures (chunks and especially knowledge structures) have been built and become efficient – this point of view is consistent with what [Kelly and Garavan](#) wrote in 2005, when they suggested that "practice on cognitive tasks...tends to be brief (hours) (in the case of novice trained), which may carry the implication that the full extent of practice related changes has not yet been examined in a cognitive task" (p. 1095) – making it possible for experts to use part of LTM as WM. The consequence of this last process is a functional reorganization.

7.2. Alternative explanation of the two patterns of neuroimaging

Before taking our explanation forward, it is important to acknowledge an alternative explanation of the two different patterns of neuroimaging results observed in the literature: the tasks used in the Neuroimaging of Experts section are of a peculiar kind compared to those used in the Neuroimaging of trained novices section. In fact, in the studies reviewed in the Neuroimaging of Experts section, knowledge is of core importance and the tasks employed are typically adapted to the domain of knowledge in order for the experts to use their knowledge. By contrast, in the Neuroimaging of trained novices section, the study tasks are standard STM/WM one's. In other words, they are very abstract tasks, built so that the participants can use their knowledge only minimally. This difference could have the consequence of making it very difficult for the participants, when trained, to build up any relevant knowledge structures, even after a very long period. This could be why no functional reorganization involving LTM areas is observed. If this explanation is correct, then there would be no need to hypothesize that the two patterns occur in a succession and that they are related to a same sequence. To rule out this possibility, we will first put forward our arguments to explain why this

explanation is not genuine and, secondly, we will explain why the two stages operate in a sequence.

Concerning the first point, it is true that the tasks used with novices and the tasks employed with experts are of a different kind. However, we feel that it is not this difference that explains the two different neuroimaging patterns, for two reasons. Firstly, it is not the case that, in the Neuroimaging of trained novices section, all the study tasks are very abstract. For example, [Jansma et al. \(2001\)](#) used letters, [Landau et al. \(2004\)](#) faces, and [Andreasen, O'Leary, Arndt, et al. \(1995\)](#) short narratives, and nevertheless brain decrease activation was observed with no functional reorganization. Therefore, we do not think that the abstractness is a key factor.

Secondly, we believe that it is always possible to develop semantic knowledge structures in WM related tasks. Let us take a spatial 3-back task as prototype of a standard task. By practicing this task, we believe that important information concerning the elements to be remembered will be picked up and encoded in an increasingly meaningful manner by the participants. Positions that share features and that are included in a sequence of 3 positions (the task is 3-back task) can potentially be chunked. Moreover, the fact that positions can share features allows categorization with high hierarchical elements containing low hierarchical elements. For example, one can imagine the following category of positions: one element from a corner and two elements from external rows. This corresponds rather directly to what [Gobet and Simon \(1996a\)](#) call a template in chess (see Section 3.1). All sequences of 3 positions that are consistent with this category are exemplars of this template. Therefore, as we wrote in Section 3.1 to characterize templates, "this means that a player will categorize different but related positions with the same template." Of course, structure needs time and practice to be cognitively extracted. And this precisely takes us to the second point, the succession of the two stages.

Alternatives to our sequential two-stage view of the two patterns of neuroimaging exist. For example, the two stages could concern two different processes operating in different situations and perhaps simultaneously. This could lead to a parallel view of the two stages. However, we do not subscribe to this point of view. First, this is because we do not think that the stages constitute two different processes. In a way, this is what the structure of the research and of the literature would make us think a priori. In fact, there is almost no link in the neuroimaging literature between the research domain where experts are studied and the research domain where novices are trained and studied. However, looking at these two separate domains through the prisms of LT-WMT and TT allows one to bridge the two patterns in a global dynamic. Concerning the view that the two stages could be two different processes, a possibility could be that the kind of tasks employed were different. However, we think that we have shown that the tasks used are not so different and that there is always the possibility to build knowledge and expertise. If the two stages are not due to two different processes, we see no reason to hypothesize a parallel view. We have suggested that they operate in sequence because templates and retrieval structure (see Section 3) are highly hierarchical structures that can incorporate chunks. They constitute a phase of expertise that is superior to the use of simple chunks. And acquiring expertise takes time. It is almost impossible to imagine retrieval structures to be built in few hours, while chunks can be built rapidly (e.g., [Cowan, Chen, & Rouder, 2004](#)). This is confirmed in terms of neuroimaging since no functional reorganization has been observed in less than 5 weeks. That is why we hypothesize a first stage where only chunks are built and used, which necessitates only a low level of expertise, and a second stage that occurs only when expertise has become stronger, which allows building templates or retrieval structures (see for a summary of the chronology of the two-stage view, Section 10).

In the next sections, we will describe and try to give an explanation for each expertise acquisition stage. As it was the aim of this article, we will attempt to show for both stages that chunks, retrieval structures and templates provide a legitimate explanation for neuroimaging data on expertise acquisition.

8. The first stage in expertise acquisition: brain decrease activation

8.1. Explaining decreases in brain activation: from chunk creation to chunk retrieval

Among the reviewed authors who observed a decrease of brain activation, two explain the decrease via the acquisition of knowledge, appealing to the chunking theory (Jansma et al., 2001; Landau et al., 2004). We completely agree and think this theory offers a good explanation of the decrease in activation in WM-related tasks. We have introduced the chunking mechanism at the beginning of the article. We would like to go further by separating two important chunking mechanisms: (a) chunk creation – the mechanism that involves binding different elements in WM into a chunk that does not yet exist in LTM; and (b) chunk retrieval – chunking mechanism that involves the perception of a chunk in WM that already exists in LTM.

Concerning the second type of chunking mechanism, it is the one that we described at the beginning of the article; if one has to remember the letters “fbiibm” and if one knows the acronyms “FBI” and “IBM,” then it is possible to chunk the information in WM into two elements “FBI” and “IBM,” because these chunks already exist in LTM.

Chunk creation contrasts with chunk retrieval. It mainly occurs when individuals do not have a strong knowledge of the information they are processing, for example in the early stages of expertise acquisition; however, it can also take place when experts are processing new information. For example, if the letters “wdg” are presented once, and if “wd” or “dg” or “wdg” do not correspond to something known, one will certainly use three chunks in WM to encode the information: w – d – g. But if the same information is presented again, one will be able to chunk more efficiently w – d – g together in WM, and if the information is displayed several times, then it is almost sure that this chunk will become an LTM chunk. At this point, if the letters “wdg” are presented again, one will be able to directly retrieve “wdg” from LTM as one chunk and therefore use only one chunk in WM to encode the information. This means that, over several presentations, a chunk that has been created can become an LTM chunk, and that the process has shifted from chunk creation to chunk retrieval.

Since novices are not experts concerning the information they are processing, the main chunking mechanism that occurs during the early stages of expertise acquisition is chunk creation. We believe that this mechanism relies strongly on WM, thus involving prefrontal WM areas (e.g., D’Esposito et al., 1995; Postle et al., 1999; Prabhakaran et al., 2000; Ungerleider et al., 1998) and parietal WM areas (e.g., Dehaene & Cohen, 1994; Postle & D’Esposito, 1999; Rodriguez et al., 1999; Todd & Marois, 2004; for the importance of parietal areas in WM, see Cowan, 2001, 2005). As the number of chunks created grows with practice, there is a gradual shift towards chunk retrieval since increasingly more chunks (in LTM) become available. We believe that this gradual shift co-occurs with a decrease of activation in prefrontal and parietal areas as less chunk creation is needed. To be more specific, and incidentally to link the chunk creation and chunk retrieval mechanisms to other established mechanisms, we put forward two reasons that can explain why the decrease

of activity in the prefrontal and parietal areas can be linked with the decrease of chunk creation and the increase of chunk retrieval.

The first reason relates to the binding mechanism. Chunk creation clearly depends on binding, which can be defined as the process of encoding the relations among stimuli that co-occur (Cohen & Eichenbaum, 1993). It is through the binding process that the elements in WM can be bound and form a chunk. The binding process is thought to occur in WM and more precisely in the focus of attention (Oakes, Ross-Sheehy, & Luck, 2006; Wheeler & Treisman, 2002).

Several theoretical models emphasize this point. Cowan and Chen (2008) have indicated that one of the functions of the focus of attention in Cowan’s WM model (Cowan, 1999, 2001, 2005) is to allow items represented concurrently to be bound into a new structure. The same idea can also be applied to the region of direct access (that is roughly compatible with the focus of attention in Cowan’s model) in Oberauer’s WM model (Oberauer, 2002; Oberauer & Lange, 2009). And in the last version of Baddeley’s model (Baddeley, 2000; Baddeley & Wilson, 2002), the binding process is said to occur in the episodic buffer (Baddeley, 2001), a WM buffer. Concerning physiology, several studies suggest that binding occurs in prefrontal regions (Prabhakaran et al., 2000; Raffone & Wolters, 2001) and in parietal regions (Oakes et al., 2006; Robertson, Treisman, Friedman-Hill, & Grabowecky, 1997; Shafritz, Gore, & Marois, 2002; Todd & Marois, 2004).

When all these results are taken together, it is easy to understand that, as more LTM chunks become available (through chunk retrieval) during the process of acquiring expertise, the necessity and opportunity to create chunks by binding separate elements becomes quantitatively less important, and therefore the activity of regions (prefrontal and parietal areas) that undergird the process of binding (and thus chunk creation) decreases.

The second reason for explaining the decrease of activity in the prefrontal and parietal areas related to practice (observed in the studies reviewed in the “Neuroimaging of Trained Novice” section) is the variation of the size of chunks in WM. In fact, whereas the number of chunks in WM is not supposed to vary with expertise (e.g., Gobet, 1998b; Lane, Gobet, & Cheng, 2001; or just slightly, Gobet & Simon, 2000) or with practice (the constant-capacity hypothesis, see for data, Chen & Cowan, 2005; Cowan et al., 2004; Tulving & Patkau, 1962; see for a review, Cowan, 2005), the chunks in WM (and therefore in LTM) can become larger with practice (Chen & Cowan, 2005; Cowan et al., 2004) and expertise (e.g., Chase & Simon, 1973a; Gobet & Simon, 1996a, 1996c). Therefore, if one can represent the same amount of information in WM with larger chunks, then the number of chunks (or the percentage of chunks)⁷ in WM necessary to represent this amount of information decreases. This could easily explain why the brain activity in WM regions decreases. In this kind of interpretation, one supposes a direct link between the pattern of decrease of activation and the decrease in the number of chunks needed. This is for example consistent with Vogel and Machizawa (2004) data. Using event-related potentials, these authors showed that, if the number of chunks necessary to represent visual information in WM diminishes, then the amplitude of an event-related waveform in posterior parietal areas and in occipital areas also decreases.

⁷ A possibility (compatible with Cowan and colleagues’ data, Chen & Cowan, 2005; Cowan et al., 2004) when chunks become larger is that instead of using a smaller number of chunks to represent the same amount of information, the number of chunks used stays stable. What can vary (i.e., decreases) is the percentage of the same number of chunks used or the amount of activation in order to activate the same number of chunks.

Table 2
Practice-related brain activation decreases in novices during working-memory-related tasks key (ranked in order of appearance from left to right in the table).

STUDIES	PARADIGM & TASK	TRAINING REGIMEN	BEHAVIORAL RESULTS	PHYSIOLOGICAL RESULTS: ACTIVATED BRAIN AREAS																
				Frontal						Cin-gulate	Parietal					Occipi-tal	Insula	Temporal		Sub-cortical
				FG	PrG	OrG	SMA	FrS	PrS		PoG	PL	IPS	Pc	AnG			OG	TG	
Garavan et al. (2000)	S-T-S Paradigm Visuospatial delayed-match-to-sample task	Short 8 hours practice	Accuracy stability and reaction times decrease	Exp.2 Mid.Bi	Exp.2 L		Exp.2 Sup.R	Exp.2 L		Exp.2 Inf.L		Exp.2 Bi		Exp.2 Mid.Bi Sup.L	Exp.2 Bi					
Jansma et al. (2001)	S-T-S Paradigm Verbal Sternberg working memory task	Short 45 min of unscanned practice and 48 min inside the scanner	Accuracy and reaction times were improved during the practiced task	inf.Bi Mid.R		L			Incr. ^a R		Sup.R									
Beauchamp et al. (2003)	S-T-S Paradigm Tower of London	Short 2 hours of practice	Decrease of time spent to solve problems (planning, execution)	Inf.L Mid.L		L	L				Sup.L Inf.L		L					CB.Bi		
Landau et al. (2004)	S-T-S Paradigm Delayed-match-to-sample task with faces	Short 30 min practice out of scanning and 48 min in the scanner	No behavioral improvements			R					Sup.L			Inf.L Mid.R	L			Th.Bi		
Sayala et al. (2006)	S-T-S Paradigm Delayed-match-to-sample task using spatial or object instructions, with faces as stimuli	Short 30 min of unscanned practice and less than one hour of practice during scanning	No behavioral improvements	Enc: Inf.Bi Mid.Bi		Enc: Bi	Enc: Sup.Bi	Enc: R	Enc: Ant.		Enc: Sup.Bi Inf.Bi		Enc: R	Enc: Mid.R	Enc: Bi		Enc: Bi	Enc: Th Cd.Bi		
Landau et al. (2007)	S-T-S Paradigm Delayed-match-to-sample task with irregular polygons (object or spatial instructions)	Short 30 min practice out of scanning and 40 min of practice during scanning	Accuracy stability and reaction times decrease	Enc: Sup.R Mid.Bi Inf.Bi	Enc: Bi			Enc: Ant.R		Enc: Sup.R Inf.R							Enc: Inf.R			
Hempel et al. (2004)	S-T-S Paradigm Spatial n-back task	Medium-length 4 weeks of practice	1) Improvement (0-2 weeks) and 2) stability of the performance (2-4 weeks)	Inf.R ^a Med.R ^a							Sup.R ^a	R ^a								
Olesen et al. (2004)	S-T-S Paradigm Visuo-spatial working memory task	Medium-length Experiment 1: 3-4 weeks practice Experiment 2: 5 weeks practice	Increase of WM span	Incr. ^b Exp.1 Mid.R			Inf.R	Sulcus	L	Incr. ^b Exp.1 Inf.R	Incr. ^b Exp.1 Bi							Incr. ^b		
Dahlin et al. (2008)	S-T-S Paradigm Updating task using letters	Medium-length 5 weeks of practice	Increase of performance in the updating task compared to a control group	Mid.R							Sup.R				Incr. ^b R	Incr. ^b R		Pa.L Pu.R		
Andreasen et al. (1995a)	T-S Paradigm Recall of a short narrative	Medium-length 1 week practice. Two sessions: 1 week and 1 day before scanning	The practiced narrative was recalled better than the unpracticed narrative			L			Ant.									CB.Bi		
Andreasen et al. (1995b)	T-S Paradigm Recall of a 15 words list	Medium-length 1 week practice. Two sessions: 1 week and 1 day before scanning	The practiced list of words was recalled better than the unpracticed list of words	Inf.Bi Sup.L							R							CB.R		
Moore et al. (2006)	T-S Paradigm Match-to-sample task with complex visual objects (polygons)	Medium-length 10 days In total 10,5 hours of training	With practiced objects, match-to-sample scores, were higher than with unpracticed objects	Incr. Inf.Bi Mid.Bi	Incr. Bi				Incr. Ant. Pos.		Incr. Bi			Incr. Mid.bi				Incr. Th.R		

S-T-S paradigm: the study used the sequence Scanning-Training-Scanning, T-S paradigm: the study used the sequence Training-Scanning, Exp: experiment, Enc: brain activation during the encoding phase, Del: brain activation during the delay phase, Rec: brain activation during the recuperation phase, Incr: increase of brain activation, R: right, L: left, Bi: bilateral, Inf: inferior, Sup: superior, Mid: middle, Ant: anterior, Pos: posterior, FG: Frontal Gyrus, PrG: Precentral Gyrus, OrG: Orbital Gyrus, SMA: Supplementary Motor Area or Pre-Supplementary Motor Area, FrS: Frontal Sulcus, PrS: Precentral Sulcus, PoG: Postcentral Gyrus, PL: Parietal Lobule, IPS: Intra-Parietal Sulcus, Pc: Precuneus, AnG: Angular Gyrus, OG: Occipital Gyrus, CalS: Calcarine Sulcus, Ins: Insula, TG: Temporal Gyrus, FuG: Fusiform Gyrus, Cd: Caudate, Th: Thalamus, CB: Cerebellum, Pu: Putamen.

^aIn these areas, there was an increase of activation in the first 2 weeks and a decrease of activation in the last 2 weeks.

^bWhen a decrease of activation was reported we have just indicated the brain areas activated, when an increase of activation was observed, we have reported the data and indicated that there was an increase.

Even if we believe that, in the activation decreasing pattern (in prefrontal and parietal areas) observed in the studies we reviewed, there is a decreasing activation that directly concerns the number of chunks, it is also likely that the link between the decrease of the number of chunks and the decrease of activation is done through other factors, such as cognitive resources. In fact, as the number of chunks in WM needed to represent the processed information in a task diminishes (because the chunks get bigger), the cognitive resources needed to execute the task also decrease; this could also explain part of the decreasing pattern in prefrontal and parietal areas.

8.2. Explaining decreases in brain activation: from chunk retrieval to knowledge structures

As mentioned above, chunks in WM (and therefore in LTM) get larger with practice and expertise. TT clearly states it: the chunks get also richer because more LTM knowledge is associated with each one of them. Moreover, several LTM chunks can become linked to knowledge. And eventually, if an individual becomes an expert, the presence of these links between several chunks can result in the creation of high-level hierarchical chunks (i.e., Ericsson and Kintsch's retrieval structures and Gobet and Simon's

templates) that we have called knowledge structures, which have large amounts of knowledge associated with them. For example, in the game of chess, templates can link to “. . . plans, moves, strategic and tactical concepts, as well as other templates” (Gobet & Simon, 1998, p. 229). We suggest that the functional reorganization of the brain can be detected in expertise acquisition when LTM chunks and knowledge structures exist and are effective in the domain of expertise.

8.3. Explaining decreases in brain activation: complementary and alternative views

In the preceding section, we have developed the idea of a chunking mechanism that would explain decrease in brain activation. We have done so because LT-WMT and TT assume the existence of these structures and these structures explain the empirical data. However, in no case do we want to deny other possible explanations.

Two types of non-exclusive explanations can be described: biological and cognitive. Concerning the first kind, several kinds of explanations have been suggested to explain decrease in brain activation. Kelly and Garavan (2005) tried to present an exhaustive list of neural mechanisms: the strengthening of synapses, “unmasking” of existing lateral connections via a modification in the local balance of excitation and inhibition, formation of new synapses, changes in neuronal processes such as dendritic length or spine density, myelination, or a combination of these (for more details, see Fields, 2011; Kolb & Gibb, 2002; Munte, Altenmüller, & Jancke, 2002; Rioult-Pedotti, Friedman, Hess, & Donoghue, 1998; Sanes & Donoghue, 2000; Schlaug, 2001). Some of these biological mechanisms are certainly the cause of decrease in brain activation. However, we believe that, in turn, these biological mechanisms are often the consequence of the chunking process described earlier.

Concerning the cognitive explanations, a plausible alternative explanation of the decrease of activation is that the task is becoming easier after practice (e.g., due to automatized response mapping, getting used to the timing of stimuli, or improved processing efficiency) and that the subjective working memory load therefore is reduced, and so are activity in load-dependent areas. At first glance, this idea seems inconsistent with our explanation. To wit, decreases in activity following practice are found in many different cognitive tasks, including simple two-choice decision tasks where chunking in working memory is not supposed to play much of a role (e.g., Kühn et al. (2011); for review, see Kelly & Garavan, 2005). We cannot rule this possibility out. However, we want to make two remarks. First, this idea is not entirely inconsistent with the chunking process, since it is possible that the task becomes easier because of this process itself. Second, if the decrease of activation is due to the task becoming easier, then one would expect that each time that there is decrease of activation there also should be an increase of performance, since with the training the task is becoming easier, all other things being equal. But this is not exactly what our review shows, since two studies did not exhibit efficiency increase while exhibiting decrease in brain activation (Landau et al., 2004; Sayala et al., 2006).

9. The second stage in expertise acquisition: brain functional reorganization

9.1. Establishing true brain functional reorganization

The functional reorganization of brain activity is considered to constitute a combined pattern of activation increases and decreases across brain areas (Kelly & Garavan, 2005). At present,

two types of functional reorganization have been pinpointed: “scaffolding” (Petersen, van Mier, Fiez, & Raichle, 1998; also labeled “redistribution” and “pseudo-reorganization” by Kelly and Garavan (2005)), and “true reorganization” (Kelly & Garavan, 2005).

According to Kelly and Garavan (2005) two criteria can be used to distinguish between scaffolding and true reorganization. The first one is the “neurophysiological pruning” of control and attentional areas. Unlike in true reorganization, scaffolding areas will be activated only at the beginning of practice and then fade away with practice. The second criterion, which seems linked to the first one, is the persistence of cognitive processes at the beginning and end of practice. In the case of scaffolding, the areas activated at the beginning are the same as those activated at the end; it is only the levels of activation of those areas that have changed. But, if “the cognitive processes are changed by practice such that practiced task performance amounts to performance of a cognitively different task, this is likely to be reflected in a neurobiologically different task map and therefore a true reorganization of functional activations” (p. 1091). We agree with Kelly and Garavan’s (2005) dichotomy about scaffolding and true functional reorganization.

Kelly and Garavan (2005) discuss several studies in light of the scaffolding vs. true functional reorganization distinction. Their review, which exclusively encompasses neuroimaging studies of trained novices, did not identify a single study showing a true functional reorganization with a WM-related task. In the “Neuroimaging of trained novices” section, we carried out a similar review, with the addition of four new studies (Dahlin et al., 2008; Landau et al., 2007; Moore et al., 2006; and Sayala et al., 2006) and the introduction of the concepts of S–T–S and S–T paradigms. Our review found evidence of one study that could be classified as true functional reorganization, Dahlin et al.’s (2008). Moreover, in our “Neuroimaging of experts” section, where we reviewed studies comparing experts and novices, we highlighted the fact that, although experts and non-experts were executing the same task, it was almost as if the two groups were “neurobiologically and cognitively” performing different tasks. This is compatible with Kelly and Garavan’s definition of true reorganization. Therefore, we consider that the type of brain functional reorganization that explains the difference between experts and novices in the studies presented in the “Neuroimaging of experts” section is of that nature. Moreover, we believe that the “true reorganization” that is involved in those studies is – as predicted by LT-WMT and TT – driven by the involvement of LTM areas, which grows with the increase of knowledge and expertise in a given task.

Below, we describe how Ericsson and Kintsch’s (1995) and Gobet and Simon’s (1996a) knowledge structures can account for the link between expertise and functional brain reorganization. It is to note that among the papers we reviewed in the “Neuroimaging of Experts” section, four appealed directly to LT-WMT or TT in order to explain their data (Campitelli et al., 2007; Pesenti et al., 2001; Saariluoma et al., 2004; Tanaka et al., 2002).

9.2. Brain functional reorganization through knowledge acquisition: the development of knowledge structures

Ericsson and Kintsch’s (1995) and Gobet and Simon’s (1996a) theories predict that the status of expert is reached only when knowledge structures – retrieval structures and templates – are fully developed, allowing experts to efficiently use part of LTM as WM. We now present more accurately how these knowledge structures account for the involvement of LTM during WM-related tasks, which leads to a functional brain reorganization in the case of experts.

With expertise, the involvement of LTM occurs in two ways: (a) in terms of semantic memory knowledge used to encode the

incoming information, and (b) in terms of transfer of the incoming information into episodic memory.

LTM intervenes through the association of semantic memory knowledge with the incoming information that is processed in WM. In Ericsson and Kintsch's theory, this is done via retrieval structures. The incoming information is associated with a set of "retrieval" LTM cues that are linked together and that constitute a retrieval structure. In the case of TT, the incoming information is associated with an undivided LTM structure (as opposed to a set of cues that are linked together), which has slots (instead of cues) that can be filled (or associated) with the aforementioned incoming information. But the important point is that both theories propose that, once individuals have become experts, and assuming that they are processing information in their area of expertise, the incoming information can be rapidly linked to knowledge structures, either retrieval structures or templates. These two knowledge structures can be conceived of as LTM structures within which the incoming information can be embedded, providing a knowledge frame for the incoming information. This is done in a way comparable to integrative schemas, scripts or frames (Bartlett, 1932; Minsky, 1975; Rumelhart, 1975; Schank & Abelson, 1977), since in both cases the incoming information can be reorganized according to familiar patterns. However, knowledge structures differ from the latter concepts at least concerning two fundamental points, which can be considered the central purposes of LT-WMT and TT. The first difference concerns their capacity to describe how, in the case of expertise, individuals can use over-practiced schemas to link and retrieve rapidly information from LTM. This may occur during WM tasks, giving the impression that WM has been expanded. We believe that this constitutes the first reason why, when doing a WM-related task, experts also activate LTM areas. The second difference, which heavily depends on the first, concerns the fast and reliable transfer of information into LTM. Due to knowledge structures, the information no longer needs to be held in WM because it has been solidly linked with semantic LTM structures and stored as an episode in episodic memory. To retrieve the information that was in WM, experts only need to activate the knowledge structures that have been associated (via cues or slots depending on the theory) with the incoming information to form an episode. Thus, when doing a WM-related task, experts also activate LTM areas.

10. General summary

The aim of this article was to try to make the PET and fMRI neuroimaging data on the role of expertise and practice in WM-related tasks more coherent by linking the psychological and physiological bases of expertise acquisition. As we have seen, results from neuroimaging studies that concern expertise in WM-related tasks are of two kinds: (a) studies using novices and a training program mostly show a decrease of cerebral activity in WM areas after a period of training and no cerebral functional reorganization, while (b) studies using trained experts tend to show results compatible with a cerebral functional reorganization. We have suggested that these two patterns of results are two stages of the same process: expertise acquisition. We believe that these two stages can be linked to two crucial psychological constructs to explain expertise: chunks and knowledge structures (Ericsson & Kintsch's, 1995, retrieval structures, and Gobet & Simon's, 1996a, templates).

We have also suggested that the discrepancy between these two patterns of results could be due to the lengths and/or the content of training periods that do not allow the trained novices to generate LTM chunks and especially knowledge structures. In line with LT-WMT (Ericsson & Kintsch, 1995) and TT (Gobet & Simon, 1996a), it is only when sufficient LTM chunks and knowledge

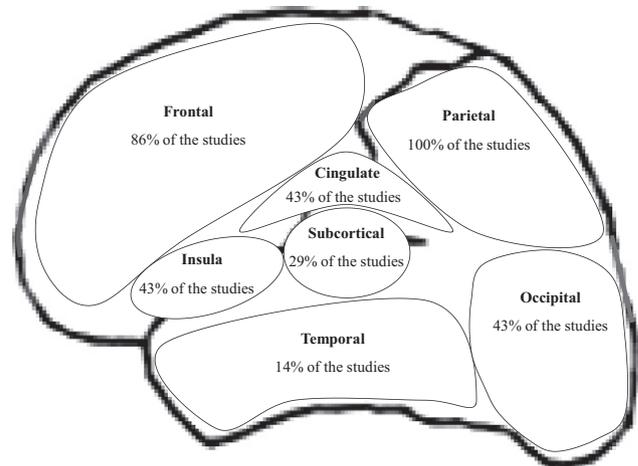


Fig. 3. Percentages of studies among those reviewed in the "Neuroimaging of trained novices" section showing a decrease of brain activation, as a function of novices' brain regions. Seven studies were included. Olesen et al. (2004) and Dahlin et al. (2008) were not included because the authors used a peculiar methodology (see Section 6.2. for comments). Andreasen, O'Leary, Arndt, et al.'s (1995), Andreasen, O'Leary, Cizadlo, et al.'s (1995) and Moore et al.'s (2006) studies were not included neither, because they used the T-S paradigm which makes difficult the comparison with studies using the S-T-S paradigm (see Section 6.3. for more precisions).

structures have been built that the second physiological stage of expertise acquisition can be observed via neuroimaging: functional brain reorganization with a strong involvement of LTM areas.

Building on these two theories, we proposed that the expertise acquisition process in WM-related tasks can be viewed as a two-physiological-stage process. At the beginning, individuals start chunking in WM, that is, bigger pieces of information are built by grouping elements of the incoming information. This has for consequence a decrease of activation (for similar suggestions in the studies reviewed, see Jansma et al., 2001; Landau et al., 2004), mainly in prefrontal and parietal areas, as shown by our "Neuroimaging of trained novices" review. Fig. 3 presents, for each specific brain region, the percentages of studies among those reviewed in the "Neuroimaging of trained novices" section showing a decrease of activation: frontal and parietal areas clearly appear among the areas exhibiting more often a decrease.

Subsequently, as practice continues, the elements that have been bound gradually begin to stabilize in LTM, first as chunks

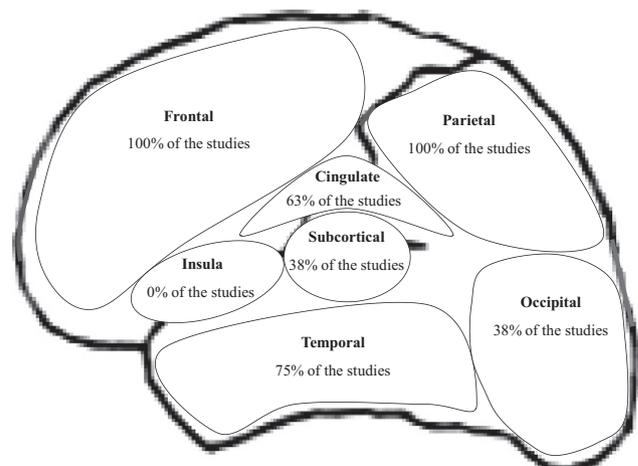


Fig. 4. Percentages of studies among those reviewed in the section "Neuroimaging of Experts" showing brain activation, as a function of experts' brain regions. All studies were included.

and then as high hierarchical chunks: knowledge structures. This has for psychological consequence an expansion of the WM capacity (see the data on expert capacities) and for physiological consequence a strong involvement of LTM areas. Fig. 4 seems consistent with this description; it shows that among the studies reviewed in the “Neuroimaging of Experts” section, the temporal areas appear to be activated very often by the experts, in addition to frontal and parietal areas, which is understandable since the tasks are WM-related. Moreover, when comparing Figs. 3 and 4, a major change between experts and novices seems to be the superior involvement – in the case of experts – of temporal regions, which are crucial for LTM.

With Fig. 5, we have tried to give a synoptic view of the dynamics of expertise acquisition, linking chunk creation, chunk retrieval and knowledge structures to the two above-mentioned physiological stages. We have used an ordinal scale-based representation (OSBR in Fig. 5), which means, for example, that at the end of the practice time, when an individual has reached a high level of expertise, chunk retrieval is more involved in expertise acquisition than chunk creation is, but less involved than knowledge structure retrieval.

In Fig. 5, we also suggest that the brain functional reorganization involving LTM areas may start when LTM chunks begin to appear in LTM, allowing chunk retrieval. However, before this can pass the statistical threshold and be detected, a vast amount of knowledge must be gathered in terms of LTM chunks and knowledge structures, even if reorganization can sometimes be detected over medium-length period, such as with the increase of activation in the temporal lobe in Dahlin et al. (2008) study. However, the minimum period of time before observing a brain functional reorganization with experts seems to be around 6 years (see the adult expert in Pesenti et al., 2001, and the child experts in Chen, Hu et al., 2006), which seems to be a more robust and conservative limit to observe functional reorganization.

Finally we would like to add that we believe that the sequence going from a decrease of brain activation to a functional reorganization involving LTM areas is the road to expertise that all individuals take when increasing their knowledge towards excellence in a domain. Even if inter-individual differences exist, the sequence is general. Certainly, differences exist in terms of semantic areas activated

in the functional reorganization (depending on the type of information manipulated in the task and the type of expertise). However, we do not think that there would be differences depending on the kind of modality of expertise. As shown by Landau et al. (2007), primary areas are less inclined to exhibit practice-related effect compared to unimodal areas and even more to multimodal ones.

The functional reorganization can be seen as an indicator of a significant amount of knowledge in a domain. Yet, we do not believe that it is a direct or indirect measure of the level of competence of an individual in his domain of expertise. This level is of course determined partly by his knowledge (captured by the functional reorganization), but it depends also on many other elements, for example fluid intelligence in chess (for more details on this topic, see Bilalić, McLeod, & Gobet, 2007). Putatively, two chess players having the same game level but not the same knowledge of the game would also differ concerning their position on the neuroimaging sequence.

11. Concluding remarks

To the best of our knowledge, this is the first time that neuroimaging studies using trained novices and neuroimaging studies using experts have been contrasted in order to be compared, thus achieving, we believe, a more in-depth analysis. We think that it is this specific comparison that allowed us to perceive the two physiological stages (decreased brain activation and brain functional reorganization) of expertise acquisition that we have proposed. And it is the proposals (concerning chunks and knowledge structures) at the core of Ericsson and Kintsch’s (1995) and Gobet and Simon’s (1996a) theories that allowed us to coherently link the two physiological stages.

In this article we have tried to foresee, before its observation, the existence of a two-stage process of expertise acquisition in WM-related tasks. To uncover this process, proper training methods and a tailored task would be needed. Concerning the task, it should be abstract and tap as little as possible any existent knowledge the participants have. As far as the training methods are concerned, they should include two kinds of training programs: (a) oriented towards the acquisition of chunks, and (b) oriented towards the acquisition of retrieval structures/templates. Therefore,

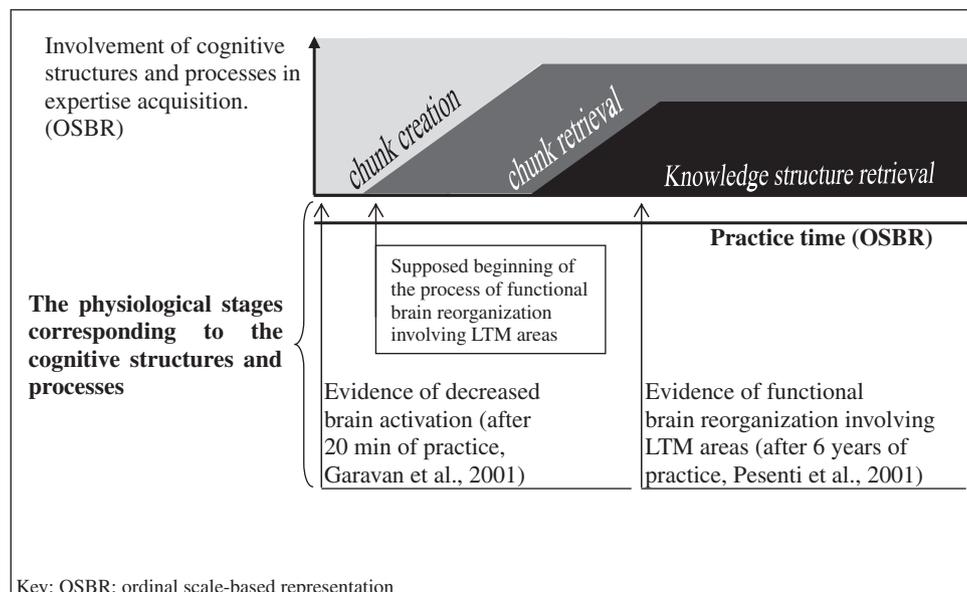


Fig. 5. Chronology of expertise acquisition in WM-related tasks, linking cognitive structures and processes to two physiological stages through LT-WMT and TT theoretical frameworks.

one could imagine three training groups. In the first group, no particular strategy would be introduced; participants would simply be executing WM tasks. In the second one, training would be built such that it favors the extraction and accumulation of knowledge concerning chunks, for example by making participants study chunks (e.g., Ericsson & Harris, 1990; Saariluoma & Laine, 2001) and by directing the participants' attention to key features of chunks. Training for the third group would be the most complex, incorporating exercises from the second training group in order to learn chunks and exercises meant to help participants to infer the hierarchical architecture of chunks by allowing them to learn and deduce families or prototypes of chunks (e.g., Gobet & Jackson, 2002). We predict that, employing specific training methods (like in the second and third group) during periods long enough (at least in order of months, see for example, Saariluoma & Laine, 2001), the results of a neuroimaging study of trained novices in WM-related tasks should be first a decrease in activation and then a cerebral functional reorganization with a gradient across the groups: the third group should evolve the fastest towards the cerebral functional reorganization pattern, while the first one should only lead to a decrease of activation.

We would like to conclude with one final remark that we think is crucial to understand the insights that were included in Ericsson and Kintsch's and Gobet and Simon's theories. We have written this paper basing ourselves on the properties of two *cognitive psychology* theories proposed in 1995 and 1996 in order to make physiological predictions about the activity of the brain when expertise is acquired. To examine, and we think to confirm, the pertinence of these predictions, we have reviewed experiments published *after* the authors' proposals and in *another field*, neuroscience. We

emphasize these differences in time and scientific fields, because it seems to us that, if a prediction describes an event that does not exist when the theory is proposed and that can be confirmed in a field for which the theory was not developed for, then the properties of the theory from which the prediction has been inferred can be considered to be "theoretically genuine."

Acknowledgments

Hubert Tardieu, our dear friend, collaborator and colleague passed away during the submission of this article. We feel privileged to have shared moments of his life. We will never forget his sense of humor, wit and genuine kindness.

We are grateful to Dr. Olivier Coubard, Pr. Pascale Piolino, Dr. Yvonnick Noël and Dr. Richard Lj Smith for comments on earlier versions of this paper and we want to deeply thank Pr. Axel Mecklinger, Dr. Martin Lövdén and one anonymous reviewer for their insightful comments.

The first author would like to personally express his profound gratitude for the inestimable help from Dr. Thierry De Greslan and the neurology service of the Val-de-Grâce hospital in Paris.

This research was supported by a grant from the French Association Nationale Pour l'Emploi (A.N.P.E.), a grant from the Association pour l'Emploi dans l'Industrie et le Commerce (A.S.S.E. D.I.C.), numéro identifiant 3048537S, and a Research Fellowship to Fernand Gobet from the Economic and Social Research Council, UK.

Appendix A

Tables 1A–4A.

Table 1A

Frequency table for BrainMap database search (October 2011) with the number of experimental comparisons identified as a function of Locations and Behavioral Domain.

		Behavioral domain in the database	
		Working memory studies	Not working memory studies
Locations in the database	Parahippocampal Gyrus and Hippocampus	72	1384
	Parahippocampal Gyrus and Hippocampus Not Activated	750	7996

Table 2A

Frequency table for BrainMap database search (October 2011) with the number of experimental comparisons identified as a function of Locations and Behavioral Domain.

		Behavioral domain in the database	
		Long-term memory explicit studies	Not long-term memory explicit studies
Locations in the database	Parahippocampal Gyrus and Hippocampus	225	1231
	Parahippocampal Gyrus and Hippocampus Not Activated	690	8056

Table 3A

Frequency table for BrainMap database search (October 2011) with the number of experimental comparisons identified as a function of Locations and Behavioral Domain.

		Behavioral domain in the database	
		Working memory studies	Not working memory studies
Locations in the database	Inferior, Middle, and Superior Frontal Gyri	577	4904
	Inferior, Middle, and Superior Frontal Gyri Not Activated	245	4476

Table 4A

Frequency table for BrainMap database search (October 2011) with the number of experimental comparisons identified as a function of Locations and Behavioral Domain.

		Behavioral domain in the database	
		Long-term memory explicit studies	Not long-term memory explicit studies
Locations in the database	Inferior, Middle, and Superior Frontal Gyri	547	4934
	Inferior, Middle, and Superior Frontal Gyri Not Activated	368	4353

References

- Aguirre, G. K., Singh, R., & D'Esposito, M. (1999). Stimulus inversion and the responses of face and object-sensitive cortical areas. *NeuroReport*, *10*, 189–194.
- Andreasen, N. C., O'Leary, D. S., Arndt, S., Cizadlo, T., Reza, K., Watkins, G. L., et al. (1995a). PET studies of memory I: Novel and practiced free recall of complex narratives. *Neuroimage*, *2*, 284–295.
- Andreasen, N. C., O'Leary, D. S., Cizadlo, T., Arndt, S., Reza, K., Watkins, G. L., et al. (1995b). PET studies of memory II: Novel versus practiced free recall of word lists. *Neuroimage*, *2*, 296–305.
- Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.
- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science*, *4*, 417–423.
- Baddeley, A. D. (2001). The magic number and the episodic buffer (Commentary on "The magical number 4 in short-term memory: A reconsideration of mental storage capacity" by N. Cowan). *Behavioural and Brain Sciences*, *24*, 117–118.
- Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews: Neuroscience*, *4*, 829–839.
- Baddeley, A. D., & Wilson, B. A. (2002). Prose recall and amnesia: Implications for the structure of working memory. *Neuropsychologia*, *40*, 1737–1743.
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.
- Beauchamp, M. H., Dagher, A., Aston, J. A., & Doyon, J. (2003). Dynamic functional changes associated with cognitive skill learning of an adapted version of the Tower of London task. *Neuroimage*, *20*, 1649–1660.
- Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, *10*, 295–307.
- Bilalić, M., McLeod, P., & Gobet, F. (2007). Does chess need intelligence? A study with young chess players. *Intelligence*, *35*, 457–470.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Rao, S. M., & Cox, R. W. (1999). Conceptual processing during the conscious resting state: A functional MRI study. *Journal of Cognitive Neuroscience*, *11*, 80–95.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *Journal of Neuroscience*, *17*, 353–362.
- Burbaud, P., Camus, O., Guehl, D., Bioulac, B., Caille, J. M., & Allard, M. (1999). A functional magnetic resonance imaging study of mental subtraction in human subjects. *Neuroscience Letters*, *273*, 195–199.
- Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, *35*, 625–641.
- Cabeza, R., Dolcos, F., Graham, R., & Nyberg, L. (2002). Similarities and differences in the neural correlates of episodic memory retrieval and working memory. *Neuroimage*, *16*, 317–330.
- Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience*, *12*, 1–47.
- Campitelli, G., Gobet, F., Head, K., Buckley, M., & Parker, A. (2007). Brain localisation of memory chunks in chessplayers. *International Journal of Neuroscience*, *117*, 1641–1659.
- Cappelletti, M., Kopelman, M. D., Morton, J., & Butterworth, B. (2005). Dissociations in numerical abilities revealed by progressive cognitive decline in a patient with semantic dementia. *Cognitive Neuropsychology*, *22*, 771–793.
- Caramazza, A., & Mahon, B. Z. (2006). The organization of conceptual knowledge in the brain: The future's past and some future directions. *Cognitive Neuropsychology*, *23*, 13–38.
- Carpenter, P. A., Just, M. A., & Reichle, E. D. (2000). Working memory and executive function: Evidence from neuroimaging. *Current Opinion in Neurobiology*, *10*, 195–199.
- Cave, C. B., & Squire, L. R. (1992). Intact verbal and nonverbal short-term memory following damage to the human hippocampus. *Hippocampus*, *2*, 151–163.
- Charness, N. (1976). Memory for chess positions: Resistance to interference. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 641–653.
- Chase, W. G., & Ericsson, K. A. (1981). Skilled memory. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 141–189). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Chase, W. G., & Simon, H. A. (1973b). Perception in chess. *Cognitive Psychology*, *4*, 55–81.
- Chase, W. G., & Simon, H. A. (1973a). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215–281). New York: Academic Press.
- Chee, M. W., O'Craven, K. M., Bergida, R., Rosen, B. R., & Savoy, R. L. (1999). Auditory and visual word processing studied with fMRI. *Human Brain Mapping*, *7*, 15–28.
- Chen, Z., & Cowan, N. (2005). Chunk limits and length limits in immediate recall: A reconciliation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 1235–1249.
- Chen, F., Hu, Z., Zhao, X., Wang, R., Yang, Z., Wang, X., et al. (2006a). Neural correlates of serial abacus mental calculation in children: A functional MRI study. *Neuroscience Letters*, *403*, 46–51.
- Chen, C. L., Wu, T. H., Cheng, M. C., Huang, Y. H., Sheu, C. Y., Hsieh, J. C., et al. (2006b). Prospective demonstration of brain plasticity after intensive abacus-based mental calculation training: An fMRI study. *Nuclear Instruments and Methods in Physics Research A*, *569*, 567–571.
- Chochon, F., Cohen, L., Van de Moortele, P. F., & Dehaene, S. (1999). Differential contributions of the left and right inferior parietal lobules to number processing. *Journal of Cognitive Neuroscience*, *11*, 617–630.
- Cloutman, L., Gottesman, R., Chaudhry, P., Davis, C., Kleinman, J. T., Pawlak, M., et al. (2009). Where (in the brain) do semantic errors come from? *Cortex*, *45*, 641–649.
- Cohen, N. J., & Eichenbaum, H. (1993). *Memory, amnesia, and the hippocampal system*. Cambridge, MA: MIT Press.
- Collette, F., Hogge, M., Salmon, E., & Van Der Linden, M. (2006). Exploration of the neural substrates of executive functioning by functional neuroimaging. *Neuroscience*, *139*, 209–221.
- Cooke, N. J., Atlas, R. S., Lane, D. M., & Berger, R. C. (1993). Role of high-level knowledge in memory for chess positions. *American Journal of Psychology*, *106*, 321–351.
- Courtney, S., Petit, L., Maisog, J., Ungerleider, L., & Haxby, J. (1998). An area specialized for spatial working memory in human frontal cortex. *Science*, *279*, 1347–1351.
- Cowan, N. (1995). *Attention and memory: An integrated framework*. London: Oxford University Press.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–185.
- Cowan, N. (2005). *Working memory capacity*. Hove, UK: Psychology Press.
- Cowan, N., Chen, Z., & Rouder, J. N. (2004). Constant capacity in an immediate serial-recall task: A logical sequel to Miller (1956). *Psychological Science*, *15*, 634–640.
- Cowan, N., & Chen, Z. (2008). How chunks form in long-term memory and affect short-term memory limits. In A. Thorn & M. Page (Eds.), *Interactions between short-term and long-term memory in the verbal domain*. Hove, East Sussex, UK: Psychology Press.
- Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 62–101). Cambridge: Cambridge University Press.
- Cowell, S. F., Egan, G. F., Code, C., Harasty, J., & Watson, J. D. (2000). The functional neuroanatomy of simple calculation and number repetition: A parametric PET activation study. *NeuroImage*, *12*, 565–573.
- Crutch, S. J., & Warrington, E. K. (2002). Preserved calculation skills in a case of semantic dementia. *Cortex*, *38*, 389–399.
- Dahlin, E., Stigsdotter-Neely, A., Larsson, A., Bäckman, L., & Nyberg, L. (2008). Transfer of learning after updating training mediated by the striatum. *Science*, *320*, 1510–1512.
- Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature*, *380*, 499–505.
- Damasio, H., Tranel, D., Grabowski, T. J., Adolphs, R., & Damasio, A. R. (2004). Neural systems behind word and concept retrieval. *Cognition*, *92*, 179–229.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466.
- de Groot, A. D. (1978). *Thought and choice in chess*. The Hague: Mouton Publishers [Original publication in 1946].
- Dehaene, S., & Cohen, L. (1994). Dissociable mechanisms of subitizing and counting: Neuropsychological evidence from simultanagnosic patients. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 958–975.
- Dehaene, S., Tzourio, N., Frak, V., Raynaud, L., Cohen, L., Mehler, J., et al. (1996). Cerebral activations during number multiplication and comparison: A PET study. *Neuropsychologia*, *34*, 1097–1106.
- Delazer, M., Karner, E., Zamarian, L., Donnemiller, E., & Benke, T. (2006). Number processing in posterior cortical atrophy: A neuropsychological case study. *Neuropsychologia*, *44*, 36–51.
- Démonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., et al. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain*, *115*, 1753–1768.
- Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *The Journal of Neuroscience*, *4*, 2051–2062.
- D'Esposito, M. (2001). Working memory. In R. Cabeza & A. Kingstone (Eds.), *Handbook of Functional Neuroimaging of Cognition* (2nd ed., pp. 292–327). Cambridge, MA: MIT Press.
- D'Esposito, M., Detre, J., Alsop, D., Shin, R., Atlas, S., & Grossman, M. (1995). The neural basis of the central executive system of working memory. *Nature*, *378*, 279–281.
- Druzgal, T. J., & D'Esposito, M. (2003). Dissecting contributions of prefrontal cortex and fusiform face area to face working memory. *Journal of Cognitive Neuroscience*, *15*, 771–784.
- Eichenbaum, H., Yonelinas, A. R., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, *30*, 123–152.
- Elliott, R., Baker, S. C., Rogers, R. D., O'Leary, D. A., Paykel, E. S., Frith, C. D., et al. (1997). Prefrontal dysfunction in depressed patients performing a complex planning task: A study using positron emission tomography. *Psychological Medicine*, *27*, 931–942.
- Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 257–297). Cambridge: Cambridge University Press.
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*, 598–601.
- Epstein, R., Stanley, D., Harris, A., & Kanwisher, N. (2000). The Parahippocampal Place Area: Perception, encoding, or memory retrieval? *Neuron*, *23*, 115–125.
- Ericsson, K. A. (1985). Memory skill. *Canadian Journal of Psychology*, *39*, 188–231.
- Ericsson, K. (2003). Exceptional memorizers: Made, not born. *Trends in Cognitive Sciences*, *7*, 233–235.

- Ericsson, K. A., & Harris, M. S. (1990). *Expert chess memory without chess knowledge*. A training study. Poster presentation at the 31st annual meeting of the psychonomics society, New Orleans.
- Ericsson, K. A., & Delaney, P. F. (1999). Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 257–297). Cambridge, MA: Cambridge University Press.
- Ericsson, K. A., Delaney, P. F., Weaver, G., & Mahadevan, R. (2004). Uncovering the structure of a memorist's superior "basic" memory capacity. *Cognitive Psychology*, 49, 191–237.
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102, 211–245.
- Ericsson, K. A., & Kintsch, W. (2000). Shortcomings of generic retrieval structures with slots of the type of Gobet (1993) proposed and modeled. *British Journal of Psychology*, 91, 571–590.
- Fields, R. D. (2011). Imaging learning: The search for a memory trace. *Neuroscientist*, 17, 185–196.
- Frey, P. W., & Adelman, P. (1976). Recall memory for visually presented chess positions. *Memory and Cognition*, 4, 541–547.
- Gabrieli, J. D., Brewer, J. B., Desmond, J. E., & Glover, G. H. (1997). Separate neural bases of two fundamental memory processes in the human medial temporal lobe. *Science*, 276, 264–266.
- Galton, C. J., Patterson, K., Graham, K., Lambon-Ralph, M. A., Williams, G., Antoun, N., et al. (2001). Differing patterns of temporal atrophy in Alzheimer's disease and semantic dementia. *Neurology*, 56, 216–225.
- Garavan, H., Kelley, D., Rosen, A., Rao, S. M., & Stein, E. A. (2000). Practice-related functional activation changes in a working memory task. *Microscopy Research and Techniques*, 51, 54–63.
- Garrard, P., & Hodges, J. R. (2000). Semantic dementia: Clinical, radiological and pathological perspectives. *Journal of Neurology*, 247, 409–422.
- Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, 3, 191–197.
- Gauthier, I., Tarr, M. J., & Anderson, A. W. (1999). Activation of the middle fusiform "face area" increases with expertise in recognizing novel objects. *Nature Neuroscience*, 2, 568–573.
- Gazzaniga, M. S., Ivry, R., & Mangun, G. R. (2009). *Cognitive neuroscience: The biology of the mind* (3rd ed.). New York: W.W. Norton.
- Gerstmann, J. (1940). Syndrome of finger agnosia, disorientation for right and left, agraphia, and acalculia. *Archives of Neurology: Psychiatry*, 44, 398–408.
- Glanzer, M., Dorfman, D., & Kaplan, B. (1981). Short-term storage in the processing of text. *Journal of Verbal Learning and Verbal Behavior*, 20, 656–670.
- Glanzer, M., Fisher, B., & Dorfman, D. (1984). Short-term storage in reading. *Journal of Verbal Learning and Verbal Behavior*, 23, 467–486.
- Gobet, F. (1998a). Expert memory: A comparison of four theories. *Cognition*, 66, 115–152.
- Gobet, F. (1998b). Memory for the meaningless: How chunks help. In *Proceedings of the twentieth meeting of the cognitive science society* (pp. 398–403). Mahwah, NJ: Erlbaum.
- Gobet, F. (2000a). Some shortcomings of long-term working memory. *British Journal of Psychology*, 91, 551–570.
- Gobet, F. (2000b). Retrieval structures and schemata: A brief reply to Ericsson and Kintsch. *British Journal of Psychology*, 91, 591–594.
- Gobet, F., & Clarkon, G. (2004). Chunks in expert memory: Evidence for the magical number four... or is it two? *Memory*, 12, 732–747.
- Gobet, F., & Jackson, S. (2002). In search of templates. *Cognitive Systems Research*, 3, 35–44.
- Gobet, F., Lane, P. C. R., Croker, S. C. H., Cheng, P., Jones, G., Oliver, I., et al. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences*, 5, 236–243.
- Gobet, F., & Simon, H. A. (1996a). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology*, 31, 1–40.
- Gobet, F., & Simon, H. A. (1996b). Recall of rapidly presented random chess positions is a function of skill. *Psychonomic Bulletin and Review*, 3, 159–163.
- Gobet, F., & Simon, H. A. (1996c). Recall of random and distorted positions: Implications for the theory of expertise. *Memory and Cognition*, 24, 493–503.
- Gobet, F., & Simon, H. A. (1998). Expert chess memory: Revisiting the chunking hypothesis. *Memory*, 6, 225–255.
- Gobet, F., & Simon, H. A. (2000). Five seconds or sixty? Presentation time in expert memory. *Cognitive Science*, 24, 651–682.
- Gorno-Tempini, M. L., Dronkers, N. F., Rankin, K. P., Ogar, J. M., Phengrasamy, L., Rosen, H. J., et al. (2004). Cognition and anatomy in three variants of primary progressive aphasia. *Annals of Neurology*, 55, 335–346.
- Gross, C. G. (1992). Representation of visual stimuli in inferior temporal cortex. *Philosophical Transactions of the Royal Society of London (Series B)*, 335, 3–10.
- Guida, A., & Tardieu, H. (2005). Is personalisation a way to operationalise long-term working memory? *Current Psychology Letters: Behaviour, Brain and Cognition*, 15, 1–17.
- Guida, A., Tardieu, H., & Nicolas, S. (2009). The personalisation method applied to a working memory task: Evidence of long-term working memory effects. *European Journal of Cognitive Psychology*, 21, 862–896.
- Hanakawa, T., Honda, M., Okada, T., Yonekura, Y., Fukuyama, H., & Shibasaki, H. (2003). Neural correlates underlying mental calculation in abacus experts: A functional magnetic resonance imaging study. *Neuroimage*, 19, 296–307.
- Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193–225). San Diego, CA: Academic Press.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, 4, 223–233.
- Hempel, A., Giesel, F. L., Caraballo, N. M. G., Amann, M., Meyer, H., Wustenberg, T., et al. (2004). Plasticity of cortical activation related to working memory during training. *American Journal of Psychiatry*, 161, 745–747.
- Hitch, G. J., Towse, J. N., & Hutton, U. (2001). What limits children's working memory span? Theoretical accounts and applications for scholastic development. *Journal of Experimental Psychology: General*, 130, 184–198.
- Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic dementia: Progressive fluent aphasia with temporal lobe atrophy. *Brain*, 115, 1783–1806.
- Hupet, M., Schelstraete, M. A., Demanet, L., & Pourtois, G. (2000). A new verbal reconstruction task to assess the integrative function of long-term working memory. *Current Psychology Letters: Behaviour, Brain and Cognition*, 2, 47–58.
- Ishai, A., Ungerleider, L. G., Martin, A., Schouten, J. L., & Haxby, J. V. (1999). Distributed representation of objects in the human ventral visual pathway. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 9379–9384.
- Jansma, J. M., Ramsey, N. F., Slagter, H. A., & Kahn, R. S. (2001). Functional anatomical correlates of controlled and automatic processing. *Journal of Cognitive Neuroscience*, 13, 730–743.
- Jeffreys, H. (1961). *Theory of probability*. New York: Oxford University Press.
- Jeneson, A., Mauldin, K. N., Hopkins, R. O., & Squire, L. R. (2011). The role of the hippocampus in retaining relational information across short delays: The importance of memory load. *Learning & Memory*, 18, 301–305.
- Jeneson, A., Mauldin, K. N., & Squire, L. R. (2010). Intact working memory for relational information after medial temporal lobe damage. *Journal of Neuroscience*, 30, 13624–13629.
- Jonides, J. (2004). How does practice makes us perfect? *Nature Neuroscience*, 7, 10–11.
- Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. S. (2008). The mind and brain of short-term memory. *Annual Review of Psychology*, 59, 193–224.
- Jonides, J., Smith, E. E., Koeppel, R. A., Awh, E., Minoshima, S., & Mintun, M. (1993). Spatial working memory as revealed by PET. *Nature*, 363, 623–625.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension. *Psychological Review*, 99, 122–149.
- Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2000). *Principles of neural science* (4th ed.). New-York: McGraw-Hill.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17, 4302–4311.
- Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition*, 68, B1–B11.
- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90, 773–795.
- Kazui, H., Kitagaki, H., & Mori, E. (2000). Cortical activation during retrieval of arithmetical facts and actual calculation: A functional magnetic resonance imaging study. *Psychiatry Clinical Neuroscience*, 54, 479–485.
- Kellogg, R. T. (2001). Long-term working memory in text production. *Memory and Cognition*, 29, 43–52.
- Kelly, A. M. C., & Garavan, H. (2005). Human functional neuroimaging of brain changes associated with practice. *Cerebral Cortex*, 15, 1089–1102.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.
- Kolb, B., & Gibb, R. (2002). Frontal lobe plasticity and behaviour. In D. Stuss & R. Knight (Eds.), *Principles of frontal lobe function* (pp. 541–556). London: Oxford University Press.
- Kühn, S., Schmiedek, F., Schott, B., Ratcliff, R., Heinze, H.-J., Düzel, E., et al. (2011). Brain areas consistently linked to individual differences in perceptual decision-making in younger as well as older adults before and after training. *Journal of Cognitive Neuroscience*, 23, 2147–2158.
- Laird, A. R., Lancaster, J. L., & Fox, P. T. (2005). BrainMap: The social evolution of a functional neuroimaging database. *Neuroinformatics*, 3, 65–78.
- Landau, S. M., Garavan, H., Schumacher, E. H., & D'Esposito, M. (2007). Regional specificity and practice: Dynamic changes in object and spatial working memory. *Brain Research*, 1180, 78–89.
- Landau, S. M., Schumacher, E. H., Garavan, H., Druzgal, T. J., & D'Esposito, M. (2004). A functional MRI study of the influence of practice on component processes of working memory. *Neuroimage*, 22, 211–221.
- Lane, P. C. R., Gobet, F., & Cheng, P. C.-H. (2001). What forms the chunks in a subject's performance? Lessons from the CHREST computational model of learning (Commentary on "The magical number 4 in short-term memory: A reconsideration of mental storage capacity" by N. Cowan). *Behavioural and Brain Sciences*, 24, 128–129.
- Lepage, M., Habib, R., & Tulving, E. (1998). Hippocampal PET activations of memory encoding and retrieval: The HIPER model. *Hippocampus*, 8, 313–322.
- Logothetis, H. K., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. *Current Biology*, 5, 552–563.
- MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, 109, 35–54.

- Maehara, Y., & Saito, S. (2007). The relationship between processing and storage in working memory span: Not two sides of the same coin. *Journal of Memory and Language*, 56, 212–228.
- Maguire, E. A., Valentine, E. R., Wilding, J. M., & Kapur, N. (2003). Routes to remembering: The brains behind superior memory. *Nature Neuroscience*, 6, 90–95.
- Martin, A. (2006). Functional neuroimaging of working memory. In R. Cabeza & A. Kingstone (Eds.), *Handbook of functional neuroimaging of cognition* (pp. 153–186). Cambridge, MA: MIT Press.
- Martin, A., & Fedio, P. (1983). Word production and comprehension in Alzheimer's disease: The breakdown of semantic knowledge. *Brain and Language*, 19, 124–141.
- Martin, A., Haxby, J., Lalonde, F., Wiggs, C., & Ungerleider, L. G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science*, 270, 102–105.
- Martin, A., Wiggs, C., Ungerleider, L., & Haxby, J. (1996). Neural correlates of category-specific knowledge. *Nature*, 379, 649–652.
- Masunaga, H., & Horn, J. L. (2000). On the emergence of wisdom: Expertise development. In W. S. Brown (Ed.), *Understanding wisdom: Sources, science, and society* (pp. 245–276). Philadelphia, PA: Templeton Foundation Press.
- McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, 9, 605–610.
- Mellet, E., Tzourio, N., Crivello, F., Joliot, M., Denis, M., & Mazoyer, B. (1996). Functional anatomy of spatial imagery generated from verbal instructions. *Journal of Neuroscience*, 16, 6504–6512.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits of our capacity for processing information. *Psychological Review*, 63, 81–97.
- Minsky, M. (1975). A framework for representing knowledge. In P. H. Winston (Ed.), *The psychology of computer vision* (pp. 211–280). New York: McGraw-Hill.
- Miyake, A. (2001). Individual differences in working memory: Introduction to the special section. *Journal of Experimental Psychology: General*, 130, 163–168.
- Moore, C., Cohen, M., & Ranganath, C. (2006). Neural mechanisms of expert skills in visual working memory. *Journal of Neuroscience*, 26, 11187–11196.
- Müller, N. G., & Knight, R. T. (2006). The functional neuroanatomy of working memory: Contributions of human brain lesion studies. *Neuroscience*, 139, 51–58.
- Mummary, C. J., Patterson, K. E., Hodges, J. R., & Price, C. J. (1998). Functional neuroanatomy of the semantic system: Divisible by what? *Journal of Cognitive Neuroscience*, 10, 766–777.
- Mummary, C. J., Patterson, K., Price, C. J., Ashburner, J., Frackowiak, R. S. J., & Hodges, J. R. (2000). A voxel-based morphometry study of semantic dementia: Relationship between temporal lobe atrophy and semantic memory. *Annals of Neurology*, 47, 36–45.
- Munte, T. F., Altenmüller, E., & Jancke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Review of Neuroscience*, 3, 473–478.
- Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology*, 7, 217–227.
- Nichols, E. A., Kao, Y.-C., Verfaellie, M., & Gabrieli, J. D. E. (2006). Working memory and long-term memory for faces: Evidence from fMRI and global amnesia for involvement of the medial temporal lobes. *Hippocampus*, 16, 604–616.
- Oakes, L. M., Ross-Sheehy, S., & Luck, S. J. (2006). Rapid development of feature binding in visual short-term memory. *Psychological Science*, 17, 781–787.
- Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 411–421.
- Oberauer, K., & Lange, E. (2009). Activation and binding in verbal working memory: A dual-process model for the recognition of nonwords. *Cognitive Psychology*, 58, 102–136.
- O'Craven, K., & Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience*, 12, 1013–1023.
- O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, 4, 95–102.
- O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford: Oxford University Press.
- Olesen, P. J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience*, 7, 75–79.
- Pesenti, M., Seron, X., Samson, D., & Duroux, B. (1999). Basic and exceptional calculation abilities in a calculating prodigy: A case study. *Mathematical Cognition*, 5, 97–148.
- Pesenti, M., Zago, L., Crivello, F., Mellet, E., Samson, D., Duroux, B., et al. (2001). Mental calculation in a prodigy is sustained by right prefrontal and medial temporal areas. *Nature Neuroscience*, 4, 103–107.
- Petersen, S. E., van Mier, H., Fiez, J. A., & Raichle, M. E. (1998). The effects of practice on the functional anatomy of task performance. *Proceedings of the National Academy of Sciences of the United States of America*, 95, 853–860.
- Poldrack, R. A. (2000). Imaging brain plasticity: Conceptual and methodological issues – A theoretical review. *NeuroImage*, 12, 1–13.
- Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, 10, 59–63.
- Postal, V. (2004). Expertise in cognitive psychology: Testing the hypothesis of long-term working memory in a study of soccer players. *Perceptual and Motor Skills*, 99, 403–420.
- Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, 139, 23–38.
- Postle, B. R., Berger, J. S., & D'Esposito, M. (1999). Functional neuroanatomical double dissociation of mnemonic and executive control processes contributing to working memory performance. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 12959–12964.
- Postle, B. R., & D'Esposito, M. (1999). "What" – Then – "Where" in visual working memory: An event-related fMRI study. *Journal of Cognitive Neuroscience*, 11, 585–597.
- Prabhakaran, V., Narayanan, K., Zhao, Z., & Gabrieli, J. D. E. (2000). Integration of diverse information in working memory within the frontal lobe. *Nature Neuroscience*, 3, 85–90.
- Raffone, A., & Wolters, G. (2001). A cortical mechanism for binding in visual working memory. *Journal of Cognitive Neuroscience*, 13, 766–785.
- Ranganath, C., & Blumenfeld, R. S. (2005). Doubts about double dissociations between short- and long-term memory. *Trends in Cognitive Science*, 9, 374–380.
- Ranganath, C., & D'Esposito, M. (2005). Directing the mind's eye: Prefrontal, inferior and medial temporal mechanisms for visual working memory. *Current Opinion in Neurobiology*, 15, 175–182.
- Rey, A. (1964). *L'examen clinique en psychologie*. Paris: Presses Universitaires de France.
- Rioult-Pedotti, M. S., Friedman, D., Hess, G., & Donoghue, J. P. (1998). Strengthening of horizontal cortical connections following skill learning. *Nature Neuroscience*, 1, 230–234.
- Robertson, L., Treisman, A., Friedman-Hill, S., & Grabowecy, M. (1997). The interaction of spatial and object pathways: Evidence from Balint's syndrome. *Journal of Cognitive Neuroscience*, 9, 295–317.
- Rodriguez, E., George, N., Lachaux, J.-P., Martinerie, J., Renault, B., & Varela, F. J. (1999). Perception's shadow: Long-distance synchronization of human brain activity. *Nature*, 397, 430–433.
- Rogers, R. D., Owen, A. M., Middleton, H. C., Williams, E. J., Pickard, J. D., Sahakian, B. J., et al. (1999). Choosing between small, likely rewards and large, unlikely rewards activates inferior and orbital prefrontal cortex. *Journal of Neuroscience*, 19, 9029–9038.
- Rowe, J., Toni, I., Josephs, O., Frackowiak, R., & Passingham, R. (2000). The prefrontal cortex: Response selection or maintenance within working memory? *Science*, 288, 1656–1660.
- Rumelhart, D. (1975). Notes on a schema for stories. In D. G. Bobrow & A. M. Collins (Eds.), *Representation and understanding: Studies in cognitive science* (pp. 211–236). New York: Academic Press.
- Rypma, B., Berger, J. S., & D'Esposito, M. (2002). The influence of working memory demand and subject performance on prefrontal cortical activity. *Journal of Cognitive Neuroscience*, 14, 721–731.
- Rypma, B., & D'Esposito, M. (1999). The roles of prefrontal brain regions in components of working memory: Effects of memory load and individual differences. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 6558–6563.
- Saariluoma, P., & Kalakoski, V. (1997). Skilled imagery and long term working memory. *American Journal of Psychology*, 110, 177–201.
- Saariluoma, P., Karlsson, H., Lyytinen, H., Teräs, M., & Geisler, F. (2004). Visuospatial representations used by chess experts: A preliminary study. *European Journal of Cognitive Psychology*, 16, 753–766.
- Saariluoma, P., & Laine, T. (2001). Novice construction of chess memory. *Scandinavian Journal of Psychology*, 42, 137–147.
- Saito, S. (2006). Memory function and short-term store as a psychological construct: Implications of a working memory framework. *Japanese Journal of Psychonomic Science*, 25, 53–60.
- Sanes, J. N., & Donoghue, J. P. (2000). Plasticity and primary motor cortex. *Annual Review of Neuroscience*, 23, 393–415.
- Sarter, M., Berntson, G. G., & Cacioppo, J. T. (1996). Brain imaging and cognitive neuroscience: Toward strong inference in attributing function to structure. *American Psychologist*, 51, 13–21.
- Sayala, S., Sala, J. B., & Courtney, S. M. (2006). Increased neural efficiency with repeated performance of a working memory task is information-type dependent. *Cerebral Cortex*, 16, 609–617.
- Schank, R., & Abelson, R. (1977). *Scripts, plans, goals and understanding*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Schlaug, G. (2001). The brain of musicians. A model for functional and structural adaptation. *Annals of the New York Academy of Sciences*, 930, 281–299.
- Schyns, P. G., Gosselin, F., & Smith, M. L. (2009). Information processing algorithms in the brain. *Trends in Cognitive Sciences*, 13, 20–26.
- Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology Neurosurgery Psychiatry*, 20, 11–21.
- Shafritz, K. M., Gore, J. C., & Marois, R. (2002). The role of the parietal cortex in visual feature binding. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 10917–10922.
- Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *Quarterly Journal of Experimental Psychology*, 22, 261–273.
- Shimamura, A. P. (1995). Memory and the frontal lobe function. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 803–813). Cambridge: MIT Press.
- Shrager, Y., Levy, D. A., Hopkins, R. O., & Squire, L. R. (2008). Working memory and the organization of brain systems. *Journal of Neuroscience*, 28, 4818–4822.
- Simon, H. A., & Gilmarin, K. J. (1973). A simulation of memory for chess positions. *Cognitive Psychology*, 5, 29–46.
- Smith, E. E., & Kosslyn, S. M. (2007). *Cognitive psychology: Mind and brain*. New Jersey: Prentice Hall.

- Sohn, Y. W., & Doane, S. M. (2003). Roles of working memory capacity and long-term working memory skill in complex task performance. *Memory and Cognition*, *31*, 458–466.
- Squire, L. R., & Alvarez, P. (1995). Retrograde amnesia and memory consolidation: A neurobiological perspective. *Current Opinion in Neurobiology*, *5*, 167–177.
- Squire, L. R., Stark, C. E. L., & Clark, R. E. (2004). The medial temporal lobe. *Annual Review of Neuroscience*, *27*, 279–306.
- Squire, L. R., & Zola-Morgan, J. (1991). The cognitive neuroscience of human memory since H.M. *Annual Review of Neuroscience*, *14*, 259–288.
- Stark, C. E. L., & Squire, L. R. (2000). fMRI activity in the medial temporal lobe during recognition memory as a function of study-test interval. *Hippocampus*, *10*, 329–337.
- Tanaka, K. (1993). Neural mechanisms of object recognition. *Science*, *262*, 685–688.
- Tanaka, S., Michimata, C., Kaminaga, T., Honda, M., & Sadato, N. (2002). Superior digit memory of abacus experts: An event-related functional MRI study. *NeuroReport*, *13*, 2187–2191.
- Tang-Wai, D. F., Graff-Radford, N. R., Boeve, B. F., Dickson, D. W., Parisi, J. E., Crook, R., et al. (2004). Clinical, genetic, and neuropathologic characteristics of posterior cortical atrophy. *Neurology*, *63*, 1168–1174.
- Tenenbaum, G., Tehan, G., Stewart, G., & Christensen, S. (1999). Recalling a floor routine: The effects of skill and age on memory for order. *Applied Cognitive Psychology*, *13*, 101–123.
- Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, *428*, 751–754.
- Towse, J. N., Hitch, G. J., & Hutton, U. (1998). A reevaluation of working memory capacity in children. *Journal of Memory and Language*, *39*, 195–217.
- Towse, J. N., Hitch, G. J., & Hutton, U. (2000). On the interpretation of working memory span in adults. *Memory and Cognition*, *28*, 341–348.
- Tulving, E., & Patkau, J. E. (1962). Concurrent effects of contextual constraint and word frequency on immediate recall and learning of verbal material. *Canadian Journal of Psychology*, *16*, 83–95.
- Tyrrell, P. J., Warrington, E. K., Frackowiak, R. S. J., & Rossor, M. N. (1990). Heterogeneity in progressive aphasia due to focal cortical atrophy: A clinical and PET scan study. *Brain*, *113*, 1321–1326.
- Ungerleider, L. G., Courtney, S. M., & Haxby, J. V. (1998). A neural system for human visual working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *95*, 883–890.
- Vandenberghe, R., Price, C. J., Wise, R., Josephs, O., & Frackowiak, R. S. J. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature*, *383*, 254–256.
- Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*, 748–751.
- Warrington, E. K. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology*, *27*, 635–657.
- Waters, G. S., & Caplan, D. (1996). The capacity theory of sentence comprehension: Critique of Just and Carpenter (1992). *Psychological Review*, *103*, 761–772.
- Wechsler, D. (1987). *Wechsler memory scale: Revised manual*. New York: Psychological Corporation.
- Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, *131*, 48–64.
- Young, B. J., Otto, T., Fox, G. D., & Eichenbaum, H. (1997). Memory representation within the parahippocampal region. *Journal of Neuroscience*, *17*, 5183–5195.
- Yovel, G., & Kanwisher, N. (2004). Face perception: Domain specific, not process specific. *Neuron*, *44*, 747–748.
- Zago, L., Pesenti, M., Mellet, E., Crivello, F., Mazoyer, B., & Tzourio-Mazoyer, N. (2001). Neural correlates of simple and complex mental calculation. *NeuroImage*, *13*, 314–327.