Journal of Expertise
2018. Vol. 1(3)
© 2018. The authors license this article under the terms of the Creative Commons Attribution 3.0 License.
ISSN 2573-2773

The Role of Long-Term Working Memory and Template Theory in Contemporary Expertise Research

Peter F. Delaney
Department of Psychology, University of North Carolina at Greensboro

Correspondence: p_delane@uncg.edu

Abstract
Long-term working memory (LT-WM) theory is a 1995 framework for understanding how experts learn skills to select relevant information, encode it into episodic long-term memory, and then retrieve it at a later time by regenerating those meaningful cues. I review the historical context in which LT-WM was proposed, including the belief that interference made long-term memory unsuitable for meeting the needs of working memory, and the changing state of working memory theories in the 1990s. I also describe a competing account called template theory that proposes that experts can rapidly store some kinds of information in episodic LTM using slotted schemas. Next, I assess some of the shortcomings of the LT-WM theory and template theory, suggesting critical areas for further research. One need is to generate adequate predictive models based on LT-WM, and to compare predictions to template theory to fully understand how working memory functions in various areas of expertise. Another is to explore how changing conceptions of working memory capacity should alter our views of the predictions of LT-WM, and to update the theory's predictions with contemporary findings.

Keywords
Expertise, long-term working memory, template theory

Introduction

More than 20 years ago, Ericsson and Kintsch (1995) proposed a major theory about how working memory is altered by expertise called *long-term working memory*, or *LT-WM*. They proposed that experts learn to use episodic long-term memory (LTM) to meet the need for temporary storage during complex processing, circumventing the usual limits of the cognitive system. LT-WM was not a new memory system, but a framework proposing that acquired skills or behaviors, when coupled with practiced knowledge, allow experts to use episodic LTM more efficiently. These skills create associative links between learned items that can later be regenerated to recover the items. Unquestionably, experts sometimes have exceptional memory for information in their domain of expertise, and LT-WM and its main competitor, template theory (e.g., Gobet & Simon, 1996), are the two major contemporary proposals for how experts expand their memory capacity.

According to PsycINFO, as of December 2018, Ericsson and Kintsch (1995) has been cited 1,037 times, and Gobet and Simon (1996) 192 times. Despite its impact, there is little

contemporary research aimed at updating original LT-WM theory, and limited attention to template theory. I will explore how LT-WM theory was constructed around theories current when it was proposed, and I will call for updates to the theory based on changes to memory theory since then. When LT-WM was proposed, storing information in episodic LTM was widely considered slow and unreliable, because researchers used individual items or unrelated pairs rather than meaningful, complex, and integrated material. To maintain rapid, reliable access, one had to use limited-capacity short-term memory (STM) to actively rehearse, postpone decay, and slowly form memories in episodic LTM. Anderson's (1980) textbook described a study (p. 170) showing that after merely 48 s, half of a list of learned paired associates was forgotten. He subsequently noted that we could sometimes maintain things longer with elaboration, but "we must be willing to make inferences and risk being wrong" (p. 200). The implied accuracy of such processes is hardly sufficient to serve as an alternative to STM.

With this understanding that episodic LTM was slow and unreliable, Chase and Simon (1973) confronted the now widely-recognized fact that real-world experts have outstanding memory for briefly-presented information. Building on an earlier study by DeGroot (1946/1978), Chase and Simon showed chess positions for 5 s to a chess beginner, a Class A player, and a master and asked them to reproduce the most recent position from memory. They obtained dramatic differences in the rate of acquisition of realistic positions, and the chess master outperformed the others, especially on midgame positions. However, for random board positions that made no sense based on the rules of chess, accuracy was virtually the same between all three players. This suggested that stronger players were better able to recognize and encode meaningful groupings of pieces in very little time. What is remarkable about this feat is not that they eventually learned the chess positions—anyone could, given enough time. What is remarkable is that the stronger players could do it so quickly. Notably, they thought it was impossible for chess masters to use episodic LTM and argued that players rehearsed pointers in STM. Each pointer specified a familiar pattern of "chunk" in LTM, which could be retrieved at the test. Later, Charness (1976) and Frey and Adesman (1976) obtained results consistent with storing chess positions in episodic LTM: They were resistant to interruptions that should disrupt rehearsal in STM, and more than one position could be retained at a time.

Another major advance was Chase and Ericsson's (1982) skilled memory theory, built to explain how people could reach extraordinary digit spans with enough practice. Verbal protocol analysis produced evidence that their digit-span experts created meaningful associations between the numbers, converting highly similar-seeming digits into meaningful configurations. Experts also constructed *retrieval structures*; that is, spatial representations that linked digit groups together. At test, experts regenerated the retrieval structure to produce retrieval cues that, combined with recency cues and meaningful associations, would help retrieve the original digits. Skilled memory theory transformed our view of episodic LTM as being slow and unreliable and suggested that unreliability occurred only when we couldn't distinguish a set of retrieval cues at the test time that could uniquely target the original memory. By building up a set of meaningful associations and guaranteed retrieval cues present at test (in the form of retrieval structures), the digit span experts could turn episodic LTM into an effective system for learning digits.

Ericsson and Kintsch's Skill-Driven Theory of LT-WM

Ericsson and Kintsch's (1995) LT-WM theory built on this foundation. It was also informed by the field's shift from a focus on STM to theories of working memory that combined processing and storage. As Shah and Miyake (1999) explained, "Everyday cognitive tasks… involve multiple steps with intermediate results that need to be kept in mind temporarily to accomplish the task at hand successfully.

'Working memory' is… the system or mechanism underlying the maintenance of task-relevant information during the performance of a cognitive task…." (p.1). While digit span experts increase *storage*, chess experts need both storage and processing (for planning and reasoning). Thus, LT-WM theory updated skilled memory theory in ways similar to how theories of working memory updated theories of STM.

Just as skilled memory theory proposed that skills mediated digit span experts' superior memory, Ericsson and Kintsch (1995) proposed that skills or behaviors could be learned to use LTM as a substitute for ordinary working memory. The skills would use relational processing to form meaningful connections between the new knowledge and pre-existing knowledge already available in LTM. Items that are strongly pre-associated in semantic memory are vastly easier to recall than less related items (e.g., Nelson, Bennett, & Leibert, 1997), so capitalizing on mediators that are highly associated can increase memory dramatically. Relational processes were sometimes supplemented with retrieval structures. Because the structures could be regenerated at test, they converted free recall into a cued recall task where the order was preserved. Furthermore, because many of the retrieval cues were unique to each representation, they sometimes allowed multiple representations of similar information to be distinguished in memory, reducing proactive and retroactive interference. LT-WM does not assume perfect memory, but it does assume that when the cues can be regenerated at test, they let people use episodic LTM to replace the usual working memory mechanisms.

To rapidly and reliably create the necessary associations during encoding, experts have to learn the necessary skills. Some of these skills deal with selecting the appropriate information to encode in the first place, which requires anticipating future uses for the information during thinking (Ericsson & Delaney, 1999). Experts also often transform what they see from perceptual categories that novices use into task-meaningful categories (e.g., "there's a pulley" vs. "it's a force problem" in physics), bringing appropriate information to bear on the solution (e.g., Chi, Feltovich, & Glaser, 1981). Likewise, even once a skill was learned, it needed to be practiced in order to speed it up to the point where it was useful as LT-WM. As an example, the memorist Rajan was able to adapt memory skills that he originally developed for digit memory to memorize lists constructed from a set of 10 arbitrary symbols (e.g., @, #, and !), reaching a symbol span of 26 after just 8 practice sessions (Ericsson, Delaney, Weaver, & Mahadevan, 2004). First, though, he had to practice recoding symbols into digits, because until he was fast enough at recoding, he couldn't capitalize on his existing skills.

LT-WM theory assumes tight memory *integration with action* and *redundancy* of encoding. Knowledge needs to be encoded in a way that supports reasoning, not just storage, which often means redundant coding since it may be used in multiple ways. That is much harder than learning how to reproduce information from memory. Consider that Ericsson and Harris (1990) taught a woman with no chess experience to memorize chess positions in 50 hours, but the resulting representations were good only for placing pieces on a memory test, not for playing the game (see also Gobet & Jackson, 2002). In contrast, chess players need to anticipate how they will use information and encode the knowledge necessary to support planning and action. Learning to do that takes much more than 50 hours, and not every domain needs to develop expert memory (for a review, see Ericsson, 2018).

Gobet and Simon's Slotted-Scheme Based Template Theory

Richman, Staszewski, and Simon (1995) proposed a competing account of expert memory specific to the digit-span expert DD in the same journal issue that LT-WM theory was proposed. Their model extended the venerable EPAM model, originally developed to explain recognition memory (Simon & Feigenbaum, 1964). EPAM uses branching decision trees called discrimination networks to learn to distinguish patterns based on their features,

learning features only when they are needed in order to correctly classify incoming stimuli. To simulate DD's memory, Richman et al. (1995) assumed knowledge was built up using discrimination networks that contained number patterns DD had learned, such as running times. For example, with experience, he might learn to distinguish "4, 6, 4, _" as a 46 minute 40-something second ten-mile time from "4, 6, _, _" as a generic 46-minute something-second ten-mile time. Second, they proposed a new mechanism whereby familiar patterns could be quickly stored into "slots" and organized using retrieval structures akin to those proposed by skilled memory theory. These slotted schemas were hierarchically-organized groups of slots that could be filled in 200 ms with items from a "type" (i.e., a restricted set of familiar items). For example, a "digit-typed" slot could only hold a digit, not a fruit; a "pattern-typed" slot could hold a pattern from the discrimination network, such as "quarter-mile time." Finally, a third kind of knowledge was production rules that represented *perceptual schemas* for properties a person might notice when studying digit lists (e.g., "the same forwards and backwards" or "two of the same pattern in a row"). EPAM-4 provided impressive quantitative predictions of time to learn a list, and both the time and accuracy in recalling a list.

Later, Gobet and Simon (1996) proposed a slotted schema-based theory called *template theory* to explain chess masters' superior chess memory. They argued that when superior memory was developed as a consequence of extended practice in thinking, and not as an end unto itself, people learn both chunks (like discrimination network patterns in Richman et al. [1995]), and slotted schemas called *templates* that are larger and contain configurations of perhaps 12 pieces, plus slots for variants. A slot could hold a piece, a square, or a familiar chunk. Consequently, one template can recognize many variants on a complex pattern. Unlike Richman et al.'s (1995) model, templates were also connected to chess playing knowledge and to one another, enabling people to use them to play chess (Campitelli & Gobet, 2005).

Template theories may provide explanations for other puzzling phenomena. Guida et al. (2018) argued that people impose an orderly structure on novel information in working memory using a template-like structure acquired while learning to read. Bilalić, McLeod, and Gobet (2008) found that chess masters' thinking could be "trapped" by the templates activated during their initial analysis of a chess position.

The major difference between LT-WM and templates is that LT-WM proposes learned *skills or behaviors* encode information rapidly in LTM, while templates propose *knowledge structures* with fillable slots (for a detailed comparison, see Ericsson & Kintsch, 2000). Even retrieval structures in LT-WM theory don't have fillable slots; they are a set of cues that get deliberately associated during study with the digits and are then regenerated later. Only when combined with other cues like recency do they uniquely specify the digit group. Both LT-WM and template theory inherit from Chase and Simon (1973) the learning of chunks of items that can serve as the basis for encoding familiar patterns.

Some recent papers have listed both possibilities without attempting to differentiate them (e.g., Guida, Gobet, Tardieu, & Nicolas, 2012). Perhaps both approaches are needed, or perhaps a new theory could replace both. Comparing LT-WM to other theories is complicated because it is a verbal framework with sometimes underspecified predictions, not a computational model (Gobet, 2000). A new generation of researchers will need to produce model-based versions of LT-WM that provide specific mechanisms, and test competing predictions derived from LT-WM and template theories. It would also be wise to integrate these future accounts with existing theories of memory, as few authors would today accept that episodic LTM is slow and unreliable for meaningful, integrated material.

The Role of Working Memory Capacity and Temporary Storage

Working memory capacity (WMC) is often measured using complex span tasks that assess the storage of information in the face of ongoing

processing. It has been linked to individual differences in controlled attention, is relatively stable over time, and predicts performance on many tasks (for a review, see Kane, Conway, Hambrick, & Engle, 2007). In 1995, we knew only that one complex span task (reading span) predicted reading comprehension (Daneman & Carpenter, 1980). Thus, it was assumed that reading relied heavily on executive resources and temporary maintenance of information during comprehension, although it was unclear whether it reflected reading-specific capacity or a general capacity (Just & Carpenter, 1992). It seemed plausible that the relationship between WMC and reading comprehension reflected merely practice at reading, as Ericsson and Kintsch (1995) argued.

Much recent work on print exposure suggests that reading experience and access to books and articles predicts reading comprehension. Stanovich, West, and Harrison (1995) showed that print exposure explained significant variance in vocabulary and general knowledge even after controlling for working memory capacity. Likewise, a meta-analysis by Mol and Bus (2011) showed that print exposure correlates reliably with reading comprehension, and its effect continue to increase with age, suggesting a reciprocal spiral. Payne, Gao, Noh, Anderson, and Stine-Morrow (2013) found that even as WMC declines with age, print exposure reduces the decline of gist memory for sentences. Hence, reading skill is clearly important to reading comprehension.

Subsequent work raised questions about the roles of WMC and LT-WM in reading, and more broadly in expertise. Post-1995 research found that complex span tasks involving little reading are strongly correlated with one another and predict reading comprehension (Kane et al., 2007), proving that reading skill alone could not account for the role of WMC in reading. More recently, McVay and Kane (2012) showed that mind wandering (which should clearly affect LT-WM encoding) fully mediated the relationship between WMC and reading skill. The result is potentially consistent with LT-WM because it suggests that for ordinary reading, as long as people are processing the text, they rapidly encode information into LT-WM. Benefits of WMC have been identified in other areas of skilled performance (e.g., Meinz & Hambrick, 2010), and these results suggest that mind wandering may be an important avenue for understanding potential roles for WMC in LT-WM.

By showing that longer attention-demanding disruptions impair comprehension accuracy, at least for low-span individuals, recent work has also challenged LT-WM theory's assertion that once a proposition is included in LT-WM, it cannot be disrupted by an interruption (Foroughi, Barragan, & Boehm-Davis, 2016; Foroughi, Werner, Barragan, & Boehm-Davis, 2015). However, Charness (1976) had long ago shown in chess, a task assumed to involve LT-WM, that there are no negative effects of most interruptions on chess reasoning; but for lengthy and attention-demanding interruptions, there were small but reliable effects (Delaney & Ericsson, 2016). Thus, although interruptions can affect LT-WM, it is the substantial *preservation* of information despite interruptions that was surprising in the first place.

Another recent development in working memory theory is the finding that WMC is related not only to attentional control, but also to long-term memory effectiveness (e.g., Daily, Lovett, & Reder, 2001; Unsworth & Engle, 2007; Unsworth, Spillers, & Brewer, 2009). These effects may be tied to the ability to use a cue to retrieve multiple associated targets (Daily et al., 2001), to use temporal and recency cues to sample representations (Unsworth & Engle, 2007), or long-lasting priming of relevant information (Was & Woltz, 2007). For ordinary representative activities such as everyday reading or selecting chess moves, recency and retrieval cues are likely to discriminate strongly active representation from prior representations. However, given the role of WMC in delimiting search sets and distinguishing representations, we might expect effects of WMC on LT-WM on (a) less-skilled individuals still developing LT-WM, or (b) when interference is maximized, such as when many chess boards were studied in a row or after extremely long interruptions.

In sum, it is worth investigating whether we need a new theory that incorporates both LT-WM or template formation and a smaller role for WMC during some tasks. We need to avoid throwing the baby out with the bathwater by assuming that any failure of memory indicates LT-WM-like mechanisms are not needed. Instead, I think we should focus our attention on delineating *when* forgetting happens from LT-WM representations and why. As in current working memory theories, either decay or interference might play a role in forgetting, and future models of LT-WM or templates will need to explore which of these is more useful for modeling experts' forgetting. Some authors have already argued that both WMC and LT-WM play a role in sentence-level decoding (e.g., Caplan & Waters, 2013), which may provide an example of how to proceed in integrating the two areas.

Author's Declaration

The author declares that there are no personal or financial conflicts of interest regarding the research in this article.

References

Anderson, J. R. (1980). *Cognitive psychology and its implications*. San Francisco, CA: W. H. Freeman.

Bilalić, M., McLeod, P., & Gobet, F. (2008). Why good thoughts block better ones: The mechanism of the pernicious Einstellung (set) effect. *Cognition, 108,* 652-661.

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

Caplan, D., & Waters, G. (2013). Memory mechanisms supporting syntactic comprehension. *Psychonomic Bulletin & Review, 20,* 243-268.

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. (1982). Skill and working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 16, pp. 1-58). New York: Academic Press.

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

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5,* 121-152.

Daily, L. Z., Lovett, M. C., & Reder, L. M. (2001). Modeling individual differences in working memory performance: A source activation account. *Cognitive Science, 25,* 315-353.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450-466.

DeGroot, A. D. (1978). *Thought and choice in chess*. The Hague: Mouton. (Original Dutch edition 1946.)

Delaney, P. F., & Ericsson, K. A. (2016). Long-term working memory and transient storage in reading comprehension: What is the evidence? *Journal of Experimental Psychology: General, 145,* 1406-1409.

Ericsson, K. A. (2018). Superior working memory in experts. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, and A. M. Williams (Eds.), *The Cambridge Handbook of Expertise and Expert Performance, 2nd Edition* (pp. 745-769). Cambridge, UK: Cambridge University Press.

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: Mechanisms of active maintenance and executive control* (pp. 257-297). Cambridge, UK: Cambridge University Press.

Ericsson, K. A., Delaney, P. F., Weaver, G. A., & Mahadevan, S. (2004). Uncovering the structure of a memorist's superior "basic" memory capacity. *Cognitive Psychology, 49,* 191-237.

Ericsson, K. A., & Harris, M. S. (1990). *Expert chess memory without chess knowledge: A training study*. Poster presented at the 31st Annual Meeting of the Psychonomic Society, New Orleans, LA.

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 that Gobet (1993) proposed and modeled. *British Journal of Psychology, 91,* 571-590.

Foroughi, C. K., Barragán, D., & Boehm-Davis, D. A.

(2016). Interrupted reading and working memory capacity. *Journal of Applied Research in Memory and Cognition, 5,* 395-400.

Foroughi, C. K., Werner, N. E., Barragán, D., & Boehm-Davis, D.A. (2015). Interruptions disrupt reading comprehension. *Journal of Experimental Psychology: General, 144,* 704–709.

Frey, P. W., & Adesman, P. (1976). Recall memory for visually presented chess positions. *Memory & Cognition, 4,* 541-547.

Gobet, F., (2000). Some shortcomings of long-term working memory. *British Journal of Psychology, 91,* 551-570.

Gobet, F., & Jackson, S. (2002). In search of templates. *Cognitive Systems Research, 3,* 35-44.

Gobet, F., & Simon, H. A. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology, 31,* 1-40.

Guida, A., Gobet, F., Tardieu, H., & Nicolas, S. (2012). How chunks, long-term working memory and templates offer a cognitive explanation for neuroimaging data on expertise acquisition: A two-stage framework. *Brain and Cognition, 79,* 221-244.

Guida, A., Megreya, A. M., Lavielle-Guida, M., Noël, Y., Mathye, F., van Dijck, J., & Abrahamse, E. (2018). Spatialization in working memory is related to literacy and reading direction: Culture "literarily" directs our thoughts. *Cognition, 175,* 96-100.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension. *Psychological Review, 99,* 122-149.

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 21-48). New York, NY: Oxford University Press.

McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology: General, 141,* 302-320.

Meinz, E. J., & Hambrick, D. Z. (2010). Deliberate practice is necessary but not sufficient to explain individual differences in piano sight-reading skill: The role of working memory capacity. *Psychological Science, 21,* 914-919.

Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin, 137,* 267-296.

Nelson, D. L., Bennett, D. J., & Leibert, T. W. (1997). One step is not enough: Making better use of association norms to predict cued recall. *Memory & Cognition, 25,* 785-796.

Payne, B. R., Gao, X., Noh, S. R., Anderson, C. J., & Stine-Morrow, E. A. L. (2013). The effects of print exposure on sentence processing and memory in older adults: Evidence for efficiency and reserve. *Neuropsychology, Development, and Cognition: Section B, Aging, Neuropsychology and Cognition, 19,* 122-149.

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

Shah, P., & Miyake, A. (1999). Models of working memory: An introduction. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 1-27). Cambridge, UK: Cambridge University Press.

Simon, H. A., & Feigenbaum, E. A. (1964). An information-processing theory of some effects of similarity, familiarization, and meaningfulness in verbal learning. *Journal of Verbal Learning and Verbal Behavior, 3,* 385-396.

Stanovich, K. E., West, R. F., & Harrison, M. R. (1995). Knowledge growth and maintenance across the life span: The role of print exposure. *Developmental Psychology, 31,* 811-826.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114,* 104-132.

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2012). Working memory capacity and retrieval limitations from long-term memory: An examination of differences in accessibility. *Quarterly Journal of Experimental Psychology, 65,* 2397-2410.

Was, C. A., & Woltz, D. J. (2007). Reexamining the relationship between working memory and comprehension: The role of available long-term memory. *Journal of Memory and Language, 56,* 86-102.

Received: 3 November 2018
Revision received: 28 December 2018
Accepted: 29 December 2018