



First published online as a Review in
Advance on September 12, 2007

The Mind and Brain of Short-Term Memory

John Jonides, Richard L. Lewis, Derek Evan Nee,
Cindy A. Lustig, Marc G. Berman,
and Katherine Sledge Moore

Department of Psychology, University of Michigan, Ann Arbor, Michigan 48109;
email: jjonides@umich.edu

Annu. Rev. Psychol. 2008.59:15.1–15.32

The *Annual Review of Psychology* is online at
<http://psych.annualreviews.org>

This article's doi:
10.1146/annurev.psych.59.103006.093615

Copyright © 2008 by Annual Reviews.
All rights reserved

0066-4308/08/0203-0001\$20.00

Key Words

working memory, attention, encoding, storage, retrieval

Abstract

The past 10 years have brought near-revolutionary changes in psychological theories about short-term memory, with similarly great advances in the neurosciences. Here, we critically examine the major psychological theories (the “mind”) of short-term memory and how they relate to evidence about underlying brain mechanisms. We focus on three features that must be addressed by any satisfactory theory of short-term memory. First, we examine the evidence for the architecture of short-term memory, with special attention to questions of capacity and how—or whether—short-term memory can be separated from long-term memory. Second, we ask how the components of that architecture enact processes of encoding, maintenance, and retrieval. Third, we describe the debate over the major reason about forgetting from short-term memory, whether interference or decay is the cause. We close with a conceptual model tracing the representation of a single item through a short-term memory task, describing the biological mechanisms that might support psychological processes on a moment-by-moment basis as an item is encoded, maintained over a delay with some forgetting, and ultimately retrieved.

Contents

INTRODUCTION.....	15.2	The Relationship of STM	
WHAT IS THE STRUCTURE		Processes to Rehearsal.....	15.14
OF SHORT-TERM MEMORY? .	15.3	WHY DO WE FORGET?	15.15
Multistore Models that		Decay Theories: Intuitive	
Differentiate Short- and		but Problematic	15.15
Long-Term Memory	15.3	Interference Theories:	
Unitary-Store Models that		Comprehensive but Complex .	15.18
Combine Short-Term and		Interference Effects	
Long-Term Memory	15.5	in Short-Term Memory.....	15.18
Controversies over Capacity	15.7	A SUMMARY OF PRINCIPLES	
Summary	15.9	AND AN ILLUSTRATION OF	
WHAT PROCESSES OPERATE		SHORT-TERM MEMORY	
ON THE STORED		AT WORK	15.20
INFORMATION?	15.10	Principles of Short-Term	
The Three Core Processes of		Memory	15.20
Short-Term Memory:		A Sketch of Short-Term Memory	
Encoding, Maintenance,		at Work.....	15.21
and Retrieval	15.10	Postscript: Revisiting Complex	
Neural Mechanisms of Short- and		Cognition.....	15.24
Long-Term Memory Retrieval .	15.13		

INTRODUCTION

Mentally add 324 and 468.
Follow the instructions to complete any
form for your federal income taxes.
Read and comprehend this sentence.

What are the features of the memory system that allows us to complete these and other complex tasks? Consider the opening example. First, you must create a temporary representation in memory for the two numbers. This representation needs to survive for several seconds to complete the task. You must then allocate your attention to different portions of the representation so that you can apply the rules of arithmetic required by the task. By one strategy, you need to focus attention on the “tens” digits (“2” and “6”) and mitigate interference from the other digits (e.g., “3” and “4”) and from partial results of previous operations (e.g., the “12” that results from adding “4” and “8”). While attending to local

portions of the problem, you must also keep accessible the parts of the problem that are not in the current focus of attention (e.g., that you now have the units digit “2” as a portion of the final answer). These tasks implicate a short-term memory (STM). In fact, there is hardly a task that can be completed without the involvement of STM, making it a critical component of cognition.

Our review relates the psychological phenomena of STM to their underlying neural mechanisms. The review is motivated by three questions that any adequate account of STM must address:

1. What is its structure?

A proper theory must describe an architecture for short-term storage. Candidate components of this architecture include storage buffers, a moving and varying focus of attention, or traces with differing levels of activation. In all cases, it is essential to provide a mechanism that allows a representation to

exist beyond the sensory stimulation that caused it or the process that retrieved the representation from long-term memory (LTM). This architecture should be clear about its psychological constructs. Furthermore, being clear about the neural mechanisms that implement those constructs will aid in development of psychological theory, as we illustrate below.

2. What processes operate on the stored information?

A proper theory must articulate the processes that create and operate on representations. Candidate processes include encoding and maintenance operations, rehearsal, shifts of attention from one part of the representation to another, and retrieval mechanisms. Some of these processes are often classified as executive functions.

3. What causes forgetting?

A complete theory of STM must account for the facts of forgetting. Traditionally, the two leading contending accounts of forgetting have relied on the concepts of decay and interference. We review the behavioral and neurophysiological evidence that has traditionally been brought to the table to distinguish decay and interference accounts, and we suggest a possible mechanism for short-term forgetting.

Most models of STM fall between two extremes: Multistore models view STM and LTM as architecturally separate systems that rely on distinct representations. By contrast, according to unitary-store models, STM and LTM rely largely on the same representations, but differ by (a) the level of activation of these representations and (b) some of the processes that normally act upon them. We focus on the distinctions drawn by these theories as we examine the evidence concerning the three questions that motivate our review. In this discussion, we assume that a representation in memory consists of a bundle of features that define a memorandum, including the context in which that memorandum was encountered.

WHAT IS THE STRUCTURE OF SHORT-TERM MEMORY?

Multistore Models that Differentiate Short- and Long-Term Memory

In his *Principles of Psychology*, William James (1890) articulated the view that short-term ("primary") memory is qualitatively different from long-term ("secondary") memory (see also Hebb 1949). The most influential successor to this view is the model of STM developed by Baddeley and colleagues (e.g., Baddeley 1986, 1992; Baddeley & Hitch 1974; Repov & Baddeley 2006). For the years 1980 to 2006, of the 16,154 papers that cited "working memory" in their titles or abstracts, fully 7339 included citations to Alan Baddeley.

According to Baddeley's model, there are separate buffers for different forms of information. These buffers, in turn, are separate from LTM. A verbal buffer, the phonological loop, is assumed to hold information that can be rehearsed verbally (e.g., letters, digits). A visuospatial sketchpad is assumed to maintain visual information and can be further fractionated into visual/object and spatial stores (Repov & Baddeley 2006, Smith et al. 1995). An episodic buffer that draws on the other buffers and LTM has been added to account for the retention of multimodal information (Baddeley 2000). In addition to the storage buffers described above, a central executive is proposed to organize the interplay between the various buffers and LTM and is implicated in controlled processing.

In short, the multistore model includes several distinctions: (a) STM is distinct from LTM, (b) STM can be stratified into different informational buffers based on information type, and (c) storage and executive processes are distinguishable. Evidence in support of these claims has relied on behavioral interference studies, neuropsychological studies, and neuroimaging data.

Evidence for the distinction between short- and long-term memory. Studies of brain-injured patients who show a deficit

in STM but not LTM or vice versa lead to the implication that STM and LTM are separate systems.¹ Patients with parietal and temporal lobe damage show impaired short-term phonological capabilities but intact LTM (Shallice & Warrington 1970, Vallar & Papagno 2002). Conversely, it is often claimed that patients with medial temporal lobe (MTL) damage demonstrate impaired LTM but preserved STM (e.g., Baddeley & Warrington 1970, Scoville & Milner 1957; we reinterpret these effects below).

Neuroimaging data from healthy subjects have yielded mixed results, however. A meta-analysis comparing regions activated during verbal LTM and STM tasks indicated a great deal of overlap in neural activation for the tasks in the frontal and parietal lobes (Cabeza et al. 2002, Cabeza & Nyberg 2000). Three studies that directly compared LTM and STM in the same subjects did reveal some regions selective for each memory system (Braver et al. 2001, Cabeza et al. 2002, Talmi et al. 2005). Yet, of these studies, only one found that the MTL was uniquely activated for LTM (Talmi et al. 2005). What might account for the discrepancy between the neuropsychological and neuroimaging data?

One possibility is that neuroimaging tasks of STM often use longer retention intervals than those employed for neuropsychological tasks, making the STM tasks more similar to LTM tasks. In fact, several studies have shown that the MTL is important when retention intervals are longer than a few seconds (Buffalo et al. 1998, Cabeza et al. 2002, Holdstock et al. 1995, Owen et al. 1995). Of the studies that compared STM and LTM in the same subjects, only Talmi et al. (2005) used an STM retention interval shorter than five seconds. This study did find, in fact, that the MTL was uniquely recruited at longer retention

intervals, providing support for the earlier neuropsychological work dissociating long- and short-term memory. As we elaborate below, however, there are other possible interpretations, especially with regard to the MTL's role in memory.

Evidence for separate buffers in short-term memory. The idea that STM can be parceled into information-specific buffers first received support from a series of studies of selective interference (e.g., Brooks 1968, den Heyer & Barrett 1971). These studies relied on the logic that if two tasks use the same processing mechanisms, they should show interfering effects on one another if performed concurrently. This work showed a double dissociation: Verbal tasks interfered with verbal STM but not visual STM, and visual tasks interfered with visual STM but not verbal STM, lending support to the idea of separable memory systems (for reviews, see Baddeley 1986 and Baddeley & Hitch 1974).

The advent of neuroimaging has allowed researchers to investigate the neural correlates of the reputed separability of STM buffers. Verbal STM has been shown to rely primarily on left inferior frontal and left parietal cortices, spatial STM on right posterior dorsal frontal and right parietal cortices, and object/visual STM on left inferior frontal, left parietal, and left inferior temporal cortices (e.g., Awh et al. 1996, Jonides et al. 1993, Smith & Jonides 1997; see review by Wager & Smith 2003). Verbal STM shows a marked left hemisphere preference, whereas spatial and object STM can be distinguished mainly by a dorsal versus ventral separation in posterior cortices (consistent with Ungerleider & Haxby 1994; see Baddeley 2003 for an account of the function of these regions in the service of STM).

The more recently postulated episodic buffer arose from the need to account for interactions between STM buffers and LTM. For example, the number of words recalled in an STM experiment can be greatly increased if the words form a sentence (Baddeley et al.

¹ Another line of neural evidence about the separability of short- and long-term memory comes from electrophysiological studies of animals engaged in short-term memory tasks. We review this evidence and its interpretation in The Architecture of Unitary-Store Models section.

1987). This “chunking” together of words to increase short-term capacity relies on additional information from LTM that can be used to integrate the words (Baddeley 2000). Thus, there must be some representational space that allows for the integration of information stored in the phonological loop and LTM. This ability to integrate information from STM and LTM is relatively preserved even when one of these memory systems is damaged (Baddeley & Wilson 2002, Baddeley et al. 1987). These data provide support for an episodic buffer that is separable from other short-term buffers and from LTM (Baddeley 2000, Baddeley & Wilson 2002). Although neural evidence about the possible localization of this buffer is thin, there is some suggestion that dorsolateral prefrontal cortex plays a role (Prabhakaran et al. 2000, Zhang et al. 2004).

Evidence for separate storage and executive processes. Baddeley’s multistore model assumes that a collection of processes act upon the information stored in the various buffers. Jointly termed the “central executive,” these processes are assumed to be separate from the storage buffers and have been associated with the frontal lobes.

Both lesion and neuroimaging data support the distinction between storage and executive processes. For example, patients with frontal damage have intact STM under conditions of low distraction (D’Esposito & Postle 1999, 2000; Malmö 1942). However, when distraction is inserted during a delay interval, thereby requiring the need for executive processes to overcome interference, patients with frontal damage show significant memory deficits (D’Esposito & Postle 1999, 2000). By contrast, patients with left temporo-parietal damage show deficits in phonological storage, regardless of the effects of interference (Vallar & Baddeley 1984, Vallar & Papagno 2002).

Consistent with these patterns, a meta-analysis of 60 functional neuroimaging studies indicated that increased demand for executive processing recruits dorsolateral frontal

cortex and posterior parietal cortex (Wager & Smith 2003). By contrast, storage processes recruit predominately posterior areas in primary and secondary association cortex. These results corroborate the evidence from lesion studies and support the distinction between storage and executive processing.

Unitary-Store Models that Combine Short-Term and Long-Term Memory

The multistore models reviewed above combine assumptions about the distinction between short-term and long-term systems, the decomposition of short-term memory into information-specific buffers, and the separation of systems of storage from executive functions. We now consider unitary models that reject the first assumption concerning distinct systems.

Contesting the idea of separate long-term and short-term systems.

The key data supporting separable short-term and long-term systems come from neuropsychology. To review, the critical contrast is between patients who show severely impaired LTM with apparently normal STM (e.g., Cave & Squire 1992, Scoville & Milner 1957) and those who show impaired STM with apparently normal LTM (e.g., Shallice & Warrington 1970). However, questions have been raised about whether these neuropsychological studies do, in fact, support the claim that STM and LTM are separable. A central question is the role of the medial temporal lobe. It is well established that the MTL is critical for long-term declarative memory formation and retrieval (Gabrieli et al. 1997, Squire 1992). However, is the MTL also engaged by STM tasks? Much research with amnesic patients showing preserved STM would suggest not, but Ranganath & Blumenfeld (2005) have summarized evidence showing that MTL is engaged in short-term tasks (see also Ranganath & D’Esposito 2005 and Nichols et al. 2006).

In particular, there is growing evidence that a critical function of the MTL is to

establish representations that involve novel relations. These relations may be among features or items, or between items and their context. By this view, episodic memory is a special case of such relations (e.g., relating a list of words to the experimental context in which the list was recently presented), and the special role of the MTL concerns its binding capabilities, not the timescale on which it operates. STM that is apparently preserved in amnesic patients may thus reflect a preserved ability to maintain and retrieve information that does not require novel relations or binding, in keeping with their preserved retrieval of remote memories consolidated before the amnesia-inducing lesion.

If this view is correct, then amnesic patients should show deficits in situations that require STM for novel relations, which they do (Hannula et al. 2005, Olson et al. 2006b). They also show STM deficits for novel materials (e.g., Buffalo et al. 1998, Holdstock et al. 1995, Olson et al. 1995, 2006a). As mentioned above, electrophysiological and neuroimaging studies support the claim that the MTL is active in support of short-term memories (e.g., Miyashita & Chang 1968, Ranganath & D'Esposito 2001). Taken together, the MTL appears to operate in both STM and LTM to create novel representations, including novel bindings of items to context.

Additional evidence for the STM-LTM distinction comes from patients with perisylvian cortical lesions who are often claimed to have selective deficits in STM (e.g., Hanley et al. 1991, Warrington & Shallice 1969). However, these deficits may be substantially perceptual. For example, patients with left perisylvian damage that results in STM deficits also have deficits in phonological processing in general, which suggests a deficit that extends beyond STM *per se* (e.g., Martin 1993).

The architecture of unitary-store models.

Our review leads to the conclusion that short- and long-term memory are not architecturally separable systems—at least not in the strong

sense of distinct underlying neural systems. Instead, the evidence points to a model in which short-term memories consist of temporary activations of long-term representations. Such unitary models of memory have a long history in cognitive psychology, with early theoretical unification achieved via interference theory (Postman 1961, Underwood & Schultz 1960). Empirical support came from demonstrations that memories in both the short and long term suffered from proactive interference (e.g., Keppel & Underwood 1962).

Perhaps the first formal proposal that short-term memory consists of activated long-term representations was by Atkinson & Shiffrin (1971, but also see Hebb 1949). The idea fell somewhat out of favor during the hegemony of the Baddeley multistore model, although it was given its first detailed computational treatment by Anderson (1983). It has recently been revived and greatly developed by Cowan (1988, 1995, 2000), McElree (2001), Oberauer (2002), Verhaeghen et al. (2004), Anderson et al. (2004), and others. The key assumption is the construct of a very limited focus of attention, although as we elaborate below, there are disagreements regarding the scope of the focus.

One shared assumption of these models is that STM consists of temporary activations of LTM representations or of representations of items that were recently perceived. The models differ from one to another regarding specifics, but Cowan's model (e.g., Cowan 2000) is representative. According to this model, there is only one set of representations of familiar material—the representations in LTM. These representations can vary in strength of activation, where that strength varies as a function of such variables as recency and frequency of occurrence. Representations that have increased strength of activation are more available for retrieval in STM experiments, but they must be retrieved nonetheless to participate in cognitive action. In addition, these representations are subject to forgetting over time. A special but limited set of these

representations, however, can be within the focus of attention, where being within the focus makes these representations immediately available for cognitive processing. According to this and similar models, then, STM is functionally seen as consisting of LTM representations that are either in the focus of attention or at a heightened level of activation.

These unitary-store models suggest a different interpretation of frontal cortical involvement in STM from multistore models. Early work showing the importance of frontal cortex for STM, particularly that of Fuster and Goldman-Rakic and colleagues, was first seen as support for multistore models (e.g., Funahashi et al. 1989, Fuster 1973, Jacobsen 1936, Wilson et al. 1993). For example, single-unit activity in dorsolateral prefrontal cortex regions (principal sulcus, inferior convexity) that was selectively responsive to memoranda during the delay interval was interpreted as evidence that these regions were the storage sites for STM. However, the sustained activation of frontal cortex during the delay period does not necessarily mean that this region is a site of STM storage. Many other regions of neocortex also show activation that outlasts the physical presence of a stimulus and provides a possible neural basis for STM representations (see Postle 2006). Furthermore, increasing evidence suggests that frontal activations reflect the operation of executive processes [including those needed to keep the representations in the focus of attention; see reviews by Postle (2006), Ranganath & D'Esposito (2005), Reuter-Lorenz & Jonides (2007), and Ruchkin et al. (2003)]. Modeling work and lesion data provide further support for the idea that the representations used in both STM and LTM are stored in those regions of cortex that are involved in initial perception and encoding, and that frontal activations reflect processes involved in selecting this information for the focus of attention and keeping it there (Damasio 1989, McClelland et al. 1995).

The principle of posterior storage also allows some degree of reconciliation between multi- and unitary-store models. Posterior

regions are clearly differentiated by information type (e.g., auditory, visual, spatial), which could support the information-specific buffers postulated by multistore models. Unitary-store models focus on central capacity limits, irrespective of modality, but they do allow for separate resources (Cowan 2000) or feature components (Lange & Oberauer 2005, Oberauer & Kliegl 2006) that occur at lower levels of perception and representation. Multi- and unitary-store models thus both converge on the idea of modality-specific representations (or components of those representations) supported by distinct posterior neural systems.

Controversies over Capacity

Regardless of whether one subscribes to multi- or unitary-store models, the issue of how much information is stored in STM has long been a prominent one (Miller 1956). Multistore models explain capacity estimates largely as interplay between the speed with which information can be rehearsed and the speed with which information is forgotten (Baddeley 1986, 1992; Repovš & Baddeley 2006). Several studies have measured this limit by demonstrating that approximately two seconds worth of verbal information can be recirculated successfully (e.g., Baddeley et al. 1975).

Unitary-store models describe capacity as limited by the number of items that can be activated in LTM, which can be thought of as the bandwidth of attention. However, these models differ on what that number or bandwidth might be. Cowan (2000) suggested a limit of approximately four items, based on performance discontinuities such as errorless performance in immediate recall when the number of items is less than four, and sharp increases in errors for larger numbers. (By this view, the classic "seven plus or minus two" is an overestimate because it is based on studies that allowed participants to engage in processes of rehearsal and chunking, and reflected contributions of both the focus and LTM; see

also Waugh & Norman 1965.) At the other extreme are experimental paradigms suggesting that the focus of attention consists of a single item (Garavan 1998, McElree 2001, Verhaeghen & Basak 2007). We briefly consider some of the central issues behind current controversies concerning capacity estimates.

Behavioral and neural evidence for the magic number 4. Cowan (2000) has reviewed an impressive array of studies leading to his conclusion that the capacity limit is four items, plus or minus one (see his Table 1). Early behavioral evidence came from studies showing sharp drop-offs in performance at three or four items on short-term retrieval tasks (e.g., Sperling 1960). These experiments were vulnerable to the criticism that this limit might reflect output interference occurring during retrieval rather than an actual limit on capacity. However, additional evidence comes from change-detection and other tasks that do not require the serial recall of individual items. For example, Luck & Vogel (1997) presented subjects with 1 to 12 colored squares in an array. After a blank interval of nearly a second, another array of squares was presented, in which one square may have changed color. Subjects were to respond whether the arrays were identical. These experiments and others that avoid the confound of output-interference (e.g., Pashler 1988) likewise have yielded capacity estimates of approximately four items.

Electrophysiological and neuroimaging studies also support the idea of a four-item capacity limit. The first such report was by Vogel & Machizawa (2004), who recorded event-related potentials (ERPs) from subjects as they performed a visual change-detection task. ERP recording shortly after the onset of the retention interval in this task indicated a negative-going wave over parietal and occipital sites that persisted for the duration of the retention interval and was sensitive to the number of items held in memory. Importantly, this signal plateaued when array size reached between three and four items. The amplitude

of this activity was strongly correlated with estimates of each subject's memory capacity and was less pronounced on incorrect than correct trials, indicating that it was causally related to performance. Subsequent functional magnetic resonance imaging (fMRI) studies have observed similar load- and accuracy-dependent activations, especially in intraparietal and intraoccipital sulci (Todd & Marois 2004, 2005). These regions have been implicated by others (e.g., Yantis & Serences 2003) in the control of attentional allocation, so it seems plausible that one rate-limiting step in STM capacity has to do with the allocation of attention (Cowan 2000; McElree 1998, 2001; Oberauer 2002).

Evidence for more severe limits on focus capacity. Another set of researchers agree there is a fixed capacity, but by measuring a combination of response time and accuracy, they contend that the focus of attention is limited to just one item (e.g., Garavan 1998, McElree 2001, Verhaeghen & Basak 2007). For example, Garavan (1998) required subjects to keep two running counts in STM, one for triangles and one for squares—as shape stimuli appeared one after another in random order. Subjects controlled their own presentation rate, which allowed Garavan to measure the time spent processing each figure before moving on. He found that responses to a figure of one category (e.g., a triangle) that followed a figure from the other category (e.g., a square) were fully 500 milliseconds longer than responses to the second of two figures from the same category (e.g., a triangle followed by another triangle). These findings suggested that attention could be focused on only one internal counter in STM at a time. Switching attention from one counter to another incurred a substantial cost in time. Using a speed-accuracy tradeoff procedure, McElree (1998) came to the same conclusion that the focus of attention contained just one item. He found that the retrieval speed for the last item in a list was substantially faster than for any other item in the list, and that other

items were retrieved at comparable rates to each other even though the accuracy of retrieval for these other items varied.

Oberauer (2002) suggested a compromise solution to the “one versus four” debate. In his model, up to four items can be directly accessible, but only one of these items can be in the focus of attention. This model is similar to that of Cowan (2000), but adds the assumption that an important method of accessing short-term memories is to focus attention on one item, depending on task demands. Thus, in tasks that serially demand attention on several items (such as those of Garavan 1998 or McElree 2001), the mechanism that accomplishes this involves changes in the focus of attention among temporarily activated representations in LTM.

Alternatives to capacity limits based on number of items. Attempting to answer the question of how many items may be held in the focus implicitly assumes that items are the appropriate unit for expressing capacity limits. Some reject this basic assumption. For example, Wilken & Ma (2004) demonstrated that a signal-detection account of STM, in which STM capacity is primarily constrained by noise, better fit behavioral data than an item-based fixed-capacity model. Recent data from change-detection tasks suggest that object complexity (Eng et al. 2005) and similarity (Awh et al. 2007) play an important role in determining capacity. Xu & Chun (2006) offer neuroimaging evidence that may reconcile the item-based and complexity accounts: In a change-detection task, they found that activation of inferior intra-parietal sulcus tracked a capacity limit of four, but nearby regions were sensitive to the complexity of the memoranda, as were the behavioral results.

Other researchers disagree with fixed item-based limits because they have demonstrated that the limit is mutable. Practice may improve subjects’ ability to use processes such as chunking to allow greater functional capacities (McElree 1998, Verhaeghen et al. 2004; but see Oberauer 2006). However, this type

of flexibility appears to alter the amount of information that can be compacted into a single representation rather than the total number of representations that can be held in STM (Miller 1956). The data of Verhaeghen et al. (2004; see Figure 5 of that paper) suggest that the latter number still approximates four, consistent with Cowan’s claims.

Building on these findings, we suggest a new view of capacity. The fundamental idea that attention can be allocated to one piece of information in memory is correct, but the definition of what that one piece is needs to be clarified. It cannot be that just one item is in the focus of attention because if that were so, hardly any computation would be possible. How could one add $3+4$, for example, if at any one time, attention could be allocated only to the “3” or the “4” or the “+” operation? We propose that attention focuses on what is bound together into a single “functional context,” whether that context is defined by time, space, some other stimulus characteristic such as semantic or visual similarity or momentary task relevance. By this account, attention can be placed on the whole problem “ $3+4$,” allowing relevant computations to be made. Complexity comes into play by limiting the number of subcomponents that can be bound into one functional context.

Summary

What are we to conclude from the data concerning the structure of STM? We favor the implication that the representational bases for perception, STM, and LTM are identical. That is, the same neural representations initially activated during the encoding of a piece of information show sustained activation during STM (or retrieval from LTM into STM; Wheeler et al. 2000 and are the repository of long-term representations. Because regions of neocortex represent different sorts of information (e.g., verbal, spatial), it is reasonable to expect that STM will have an organization by type of material as well. Functionally, memory in the short term seems to consist of items

in the focus of attention along with recently attended representations in LTM. These items in the focus of attention number no more than four, and they may be limited to just a single representation (consisting of items bound within a functional context).

We turn below to processes that operate on these representations.

WHAT PROCESSES OPERATE ON THE STORED INFORMATION?

Theoretical debate about the nature of STM has been dominated by discussion of structure and capacity, but the issue of process is also important. Verbal rehearsal is perhaps most intuitively associated with STM and plays a key role in the classic model (Baddeley 1986). However, as we discuss below, rehearsal most likely reflects a complex strategy rather than a primitive STM process. Modern approaches offer a large set of candidate processes, including encoding and maintenance (Ranganath et al. 2004), attention shifts (Cowan 2000), spatial rehearsal (Awh & Jonides 2001), updating (Oberauer 2005), overwriting (Neath & Nairne 1995), cue-based parallel retrieval (McElree 2001), and interference-resolution (Jonides & Nee 2006).

Rather than navigating this complex and growing list, we take as our cornerstone the concept of a limited focus of attention. The central point of agreement for the unitary-store models discussed above is that there is a distinguishable focus of attention in which representations are directly accessible and available for cognitive action. Therefore, it is critical that all models must identify the processes that govern the transition of memory representations into and out of this focused state.

The Three Core Processes of Short-Term Memory: Encoding, Maintenance, and Retrieval

If one adopts the view that a limited focus of attention is a key feature of short-term stor-

age, then understanding processing related to this limited focus amounts to understanding three basic types of cognitive events²: (a) encoding processes that govern the transformation from perceptual representations into the cognitive/attentional focus, (b) maintenance processes that keep information in the focus (and protect it from interference or decay), and (c) retrieval processes that bring information from the past back into the cognitive focus (possibly reactivating perceptual representations).

Encoding of items into the focus. Encoding processes are the traditional domain of theories of perception and are not treated explicitly in any of the current major accounts of STM. Here we outline three implicit assumptions about encoding processes made in most accounts of STM, and we assess their empirical and theoretical support.

First, the cognitive focus is assumed to have immediate access to perceptual processing—that is, the focus may include contents from the immediate present as well as contents retrieved from the immediate past. In Cowan's (2000) review of evidence in favor of the number four in capacity estimates, several of the experimental paradigms involve focused representations of objects in the immediate perceptual present or objects presented less than a second ago. These include visual tracking experiments (Pylyshyn et al. 1994), enumeration (Trick & Pylyshyn 1993), and whole report of spatial arrays and spatiotemporal arrays (Darwin et al. 1972, Sperling 1960). Similarly, in McElree's (2006) and Garavan's (1998) experiments, each incoming item in the stream of material (words or letters or objects) is assumed to be represented momentarily in the focus.

²This carving up of STM processes is also consistent with recent approaches to individual differences in working memory, which characterize individual variation not in terms of variation in buffer capacity, but rather in variation in maintenance and retrieval processes (Unsworth & Engle 2007).

Second, all of the current theories assume that perceptual encoding into the focus of attention results in a displacement of other items from the focus. For example, in McElree's single-item focus model, each incoming item not only has its turn in the focus, but it also replaces the previous item. On the one hand, the work reviewed above regarding performance discontinuities after the putative limit of STM capacity has been reached appears to support the idea of whole-item displacement. On the other hand, as also described above, this limit may be susceptible to factors such as practice and stimulus complexity. An alternative to whole-item displacement as the basis for interference is a graded similarity-based interference, in which new items entering the focus may partially overwrite features of the old items or compete with old items to include those featural components in their representations as a function of their similarity. At some level, graded interference is clearly at work in STM, as Nairne (2002) and others have demonstrated (we review this evidence in more detail below). But the issue at hand is whether the focus is subject to such graded interference, and if such interference is the process by which encoding (or retrieving) items into the focus displaces prior items. Although there does not appear to be evidence that bears directly on this issue (the required experiments would involve manipulations of similarity in just the kinds of paradigms that Cowan, McElree, Oberauer, and others have used to provide evidence for the limited focus), the performance discontinuities strongly suggest that something like displacement is at work.

Third, all of the accounts assume that perceptual encoding does not have obligatory access to the focus. Instead, encoding into the focus is modulated by attention. This follows rather directly from the assumptions about the severe limits on focus capacity: There must be some controlled way of directing which aspects of the perceptual present, as well as the cognitive past, enter into the focused state. Stated negatively, there must be some way of

preventing aspects of the perceptual present from automatically entering into the focused state. Postle (2006) recently found that increased activity in dorsolateral prefrontal cortex during the presentation of distraction during a retention interval was accompanied by a selective decrease in inferior temporal cortical activity. This pattern suggests that prefrontal regions selectively modulated posterior perceptual areas to prevent incoming sensory input from disrupting the trace of the task-relevant memorandum.

In summary, current approaches to STM have an obligation to account for how controlled processes bring relevant aspects of perception into cognitive focus and leave others out. It is by no means certain that existing STM models and existing models of perceptual attention are entirely compatible on this issue, and this is a matter of continued lively debate (Milner 2001, Schubert & Frensch 2001, Woodman et al. 2001).

Maintenance of items in the focus. Once an item is in the focus of attention, what keeps it there? If the item is in the perceptual present, the answer is clear: attention-modulated, perceptual encoding. The more pressing question is: What keeps something in the cognitive focus when it is not currently perceived? For many neuroscientists, this is the central question of STM—how information is held in mind for the purpose of future action after the perceptual input is gone. There is now considerable evidence from primate models and from imaging studies on humans for a process of active maintenance that keeps representations alive and protects them from irrelevant incoming stimuli or intruding thoughts (e.g., Postle 2006).

We argue that this process of maintenance is not the same as rehearsal. Indeed, the number of items that can be maintained without rehearsal forms the basis of Cowan's (2000) model. Under this view, rehearsal is not a basic process but rather is a strategy for accomplishing the functional demands for sustaining memories in the short term—a

strategy composed of a series of retrievals and re-encodings. We consider rehearsal in more detail below, but we consider here the behavioral and neuroimaging evidence for maintenance processes.

There is now considerable evidence from both primate models and human electroencephalography and fMRI studies for a set of prefrontal-posterior circuits underlying active maintenance. Perhaps the most striking is the classic evidence from single-cell recordings showing that some neurons in prefrontal cortex fire selectively during the delay period in delayed-match-to-sample tasks (e.g., Funahashi et al. 1989, Fuster 1973). As mentioned above, early interpretations of these frontal activations linked them directly to STM representations (Goldman-Rakic 1987), but more recent theories suggest they are part of a frontal-posterior STM circuit that maintains representations in posterior areas (Pasternak & Greenlee 2005, Ranganath 2006, Ruchkin et al. 2003). Furthermore, as described above, maintenance operations may modulate perceptual encoding to prevent incoming perceptual stimuli from disrupting the focused representation in posterior cortex (Postle 2006). Several computational neural-network models of circuits for maintenance hypothesize that prefrontal cortical circuits support attractors, self-sustaining patterns observed in certain classes of recurrent networks (Hopfield 1982, Rougier et al. 2005, Polk et al. 2002). A major challenge is to develop computational models that are able to engage in active maintenance of representations in posterior cortex while simultaneously processing, to some degree, incoming perceptual material (see Renart et al. 1999 for a related attempt).

Retrieval of items into the focus. Many of the major existing STM architectures are silent on the issue of retrieval. However, all models that assume a limited focus also assume that there is some means by which items outside that focus (either in a dormant long-term store or in some highly activated portion

of LTM) are brought into the focus by switching the attentional focus onto those items. Following Sternberg (1966), McElree (2006), and others, we label this process “retrieval.” Despite this label, it is important to keep in mind that the associated spatial metaphor of an item moving from one location to another is misleading given our assumption about the common neural representations underlying STM and LTM.

There is now considerable evidence, mostly from mathematical models of behavioral data, that STM retrieval of item information is a rapid, parallel, content-addressable process. The current emphasis on parallel search processes is quite different from the earliest models of STM retrieval, which postulated a serial scanning process (i.e., Sternberg 1966; see McElree 2006 for a recent review and critique). Serial-scanning models fell out of favor because of empirical and modeling work showing that parallel processes provide a better account of the reaction time distributions in STM tasks (e.g., Hockley 1984). For example, McElree has created a variation on the Sternberg recognition probe task that provides direct support for parallel, rather than serial, retrieval. In the standard version of the task, participants are presented with a memory set consisting of a rapid sequence of verbal items (e.g., letters or digits), followed by a probe item. The task is to identify whether the probe was a member of the memory set. McElree & Doshier's (1989) innovation was to manipulate the deadline for responding. The time course of retrieval (accuracy as a function of response deadline) can be separately plotted for each position within the presentation sequence, allowing independent assessments of accessibility (how fast an item can be retrieved) and availability (asymptotic accuracy) as a function of set size and serial position. Many experiments yield a uniform rate of access for all items except for the most recent item, which is accessed more quickly. The uniformity of access rate is evidence for parallel access, and the distinction between the most

recent item and the other items is evidence for a distinguished focus of attention.

Neural Mechanisms of Short- and Long-Term Memory Retrieval

The cue-based retrieval processes described above for STM are very similar to those posited for LTM (e.g., Anderson et al. 2004, Gillund & Shiffrin 1984, Murdock 1982). As a result, retrieval failures resulting from similarity-based interference and cue overlap are ubiquitous in both STM and LTM. Both classic studies of recall from STM (e.g., Keppel & Underwood 1962) and more recent studies of interference in probe-recognition tasks (e.g., Jonides & Nee 2006, McElree & Doshier 1989, Monsell 1978) support the idea that interference plays a major role in forgetting over short retention intervals as well as long ones (see below). These common effects would not be expected if STM retrieval were a different process restricted to operate over a limited buffer, but they are consistent with the notion that short-term and long-term retrieval are mediated by the same cue-based mechanisms.

The heavy overlap in the neural substrates for short-term and long-term retrieval provides additional support for the idea that retrieval processes are largely the same over different retention intervals. A network of medial temporal regions, lateral prefrontal regions, and anterior prefrontal regions has been extensively studied and shown to be active in long-term retrieval tasks (e.g., Buckner et al. 1998, Cabeza & Nyberg 2000, Fletcher & Henson 2001). We reviewed above the evidence for MTL involvement in both short- and long-term memory tasks that require novel representations (see section titled "Contesting the Idea of Separate Long-Term and Short-Term Systems"). Here, we examine whether the role of frontal cortex is the same for both short- and long-term retrieval.

The conclusion derived from neuroimaging studies of various different STM procedures is that this frontal role is the

same in short-term and long-term retrieval. For example, several event-related fMRI studies of the retrieval stage of the probe-recognition task found increased activation in lateral prefrontal cortex similar to the activations seen in studies of LTM retrieval (e.g., D'Esposito et al. 1999, D'Esposito & Postle 2000, Manoach et al. 2003). Badre & Wagner (2005) also found anterior prefrontal activations that overlapped with regions implicated in episodic recollection. The relatively long retention intervals often used in event-related fMRI studies leaves them open to the criticism that by the time of the probe, the focus of attention has shifted elsewhere, causing the need to retrieve information from LTM (more on this discussion below). However, a meta-analysis of studies that involved bringing very recently presented items to the focus of attention likewise found specific involvement of lateral and anterior prefrontal cortex (Johnson et al. 2005). These regions appear to be involved in retrieval, regardless of timescale.

The same conclusion may be drawn from recent imaging studies that have directly compared long- and short-term retrieval tasks using within-subjects designs (Cabeza et al. 2002, Ranganath et al. 2003, Talami et al. 2005). Ranganath et al. (2003) found the same bilateral ventrolateral and dorsolateral prefrontal regions engaged in both short- and long-term tasks. In some cases, STM and LTM tasks involve the same regions but differ in the relative amount of activation shown within those regions. For example, Cabeza et al. (2002) reported similar engagement of medial temporal regions in both types of task, but greater anterior and ventrolateral activation in the long-term episodic tasks. Talami et al. (2005) reported greater activation in both medial temporal and lateral frontal cortices for recognition probes of items presented early in a 12-item list (presumably necessitating retrieval from LTM) versus items presented later in the list (presumably necessitating retrieval from STM). One possible reason for this discrepancy is that recognition for

late-list items did not require retrieval because these items were still in the focus of attention. This account is plausible since late-list items were drawn either from the last-presented or second-to-last presented item and preceded the probe by less than two seconds.

In summary, the bulk of the neuroimaging evidence points to the conclusion that the activation of frontal and medial temporal regions depends on whether the information is currently in or out of focus, not whether the task nominally tests STM or LTM. Similar reactivation processes occur during retrieval from LTM and from STM when the active maintenance has been interrupted (see Sakai 2003 for a more extensive review).

The Relationship of STM Processes to Rehearsal

Notably, our account of core STM processes excludes rehearsal. How does rehearsal fit in? We argue that rehearsal is simply a controlled sequence of retrievals and re-encodings of items into the focus of attention (Baddeley 1986, Cowan 1995). The theoretical force of this assumption can be appreciated by examining the predictions it makes when coupled with our other assumptions about the structures and processes of the underlying STM architecture. Below we outline these predictions and the behavioral, developmental, neuroimaging, and computational work that support this view.

Rehearsal as retrieval into the focus.

When coupled with the idea of a single-item focus, the assumption that rehearsal is a sequence of retrievals into the focus of attention makes a very clear prediction: A just-rehearsed item should display the same retrieval dynamics as a just-perceived item. McElree (2006) directly tested this prediction using a version of his response-deadline recognition task, in which subjects were given a retention interval between presentation of the list and the probe rather than presented with the probe immediately after the list. Subjects were ex-

plicitly instructed to rehearse the list during this interval and were trained to do so at a particular rate. By controlling the rate, it was possible to know when each item was rehearsed and hence re-established in the focus. The results were compelling: When an item was predicted to be in focus because it had just been rehearsed, it showed the same fast retrieval dynamics as an item that had just been perceived. In short, the speed-accuracy tradeoff functions showed the familiar in-focus/out-of-focus dichotomy of the standard paradigm, but the dichotomy was established for internally controlled rehearsal as well as externally controlled perception.

Rehearsal as strategic retrieval. Rehearsal is often implicitly assumed as a component of active maintenance, but formal theoretical considerations of STM typically take the opposite view. For example, Cowan (2000) provides evidence that although first-grade children do not use verbal rehearsal strategies, they nevertheless have measurable focus capacities. In fact, Cowan (2000) uses this evidence to argue that the performance of very young children is revealing of the fundamental capacity limits of the focus of attention because it is not confounded with rehearsal.

If rehearsal is the controlled composition of more primitive STM processes, then rehearsal should activate the same brain circuits as the primitive processes, possibly along with additional (frontal) circuits associated with their control. In other words, there should be overlap of rehearsal with brain areas subserving retrieval and initial perceptual encoding. Likewise, there should be control areas distinct from those of the primitive processes.

Both predictions receive support from neuroimaging studies. The first prediction is broadly confirmed: There is now considerable evidence for the reactivation of areas associated with initial perceptual encoding in tasks that require rehearsal (see Jonides et al. 2005 for a recent review; note also that

evidence exists for reactivation in LTM retrieval: Wheeler 2000, 2006).

The second prediction—that rehearsal engages additional control areas beyond those participating in maintenance, encoding, and retrieval—receives support from two effects. One is that verbal rehearsal engages a set of frontal structures associated with articulation and its planning: supplementary motor, premotor, inferior frontal, and posterior parietal areas (e.g., Chein & Fiez 2001, Jonides et al. 1998, Smith & Jonides 1999). The other is that spatial rehearsal engages attentionally mediated occipital regions, suggesting rehearsal processes that include retrieval of spatial information (Awh et al. 1998, 1999, 2001).

Computational modeling relevant to strategic retrieval. Finally, prominent symbolic and connectionist computational models of verbal STM tasks are based on architectures that do not include rehearsal as a primitive process, but rather assume it as a strategic composition of other processes operating over a limited focus. The Burgess & Hitch (2005, 2006) connectionist model, the Executive-Process/Interactive Control (EPIC) symbolic model (Meyer and Kieras 1997), and the Atomic Components of Thought (ACT-R) hybrid model (Anderson & Matessa 1997) all assume that rehearsal in verbal STM consists of a controlled sequence of retrievals of items into a focused state. They all assume different underlying mechanisms for the focus (the Burgess & Hitch model has a winner-take-all network; ACT-R has an architectural buffer with a capacity of one chunk; EPIC has a special auditory store), but all assume strategic use of this focus to accomplish rehearsal. These models jointly represent the most successful attempts to account for a range of detailed empirical phenomena traditionally associated with rehearsal, especially in verbal serial recall tasks. Their success therefore provides further support for the plausibility of a compositional view of rehearsal.

WHY DO WE FORGET?

Forgetting in STM is a vexing problem: What accounts for failures to retrieve something encoded just seconds ago? There are two major explanations for forgetting, often placed in opposition: time-based decay and similarity-based interference. Below, we describe some of the major findings in the literature related to each of these explanations, and we suggest that they may ultimately result from the same underlying principles.

Decay Theories: Intuitive but Problematic

The central claim of decay theory is that as time passes, information in memory erodes, and so it is less available for later retrieval. This explanation has strong intuitive appeal. However, over the years there have been sharp critiques of decay, questioning whether it plays any role at all (for recent examples, see Lewandowsky et al. 2004 and the review in this journal by Nairne 2002).

Decay explanations are controversial for two reasons: First, experiments attempting to demonstrate decay can seldom eliminate alternative explanations. For example, Keppel & Underwood (1962) demonstrated that forgetting in the classic Brown-Peterson paradigm (designed to measure time-based decay) was due largely, if not exclusively, to proactive interference from prior trials. Second, without an explanation of how decay occurs, it is difficult to see decay theories as more than a restatement of the problem. Some functional arguments have been made for the usefulness of the notion of memory decay—that decaying activations adaptively mirror the likelihood that items will need to be retrieved (Anderson & Schooler 1991), or that decay is functionally necessary to reduce interference (Altmann & Gray 2002). Nevertheless, McGeoch's famous (1932) criticism of decay theories still holds merit: Rust does not occur because of time itself, but rather from oxidation processes that occur with time. Decay theories must explain

the processes by which decay could occur, i.e., they must identify the oxidation process in STM.

Retention-interval confounds: controlling for rehearsal and retroactive interference.

The main problem in testing decay theories is controlling for what occurs during the retention interval. Many experiments include an attention-demanding task to prevent participants from using rehearsal that would presumably circumvent decay. However, a careful analysis of these studies by Roediger et al. (1977) makes one wonder whether the use of a secondary task is appropriate to prevent rehearsal at all. They compared conditions in which a retention interval was filled by nothing, by a relatively easy task, or by a relatively difficult one. Both conditions with a filled interval led to worse memory performance, but the difficulty of the intervening task had no effect. Roediger et al. (1977) concluded that the primary memory task and the interpolated task, although demanding, used different processing pools of resources, and hence the interpolated tasks may not have been effective in preventing rehearsal. So, they argued, this sort of secondary-task technique may not prevent rehearsal and may not allow for a convincing test of a decay hypothesis.

Another problem with tasks that fill the retention interval is that they require subjects to use STM (consider counting backward, as in the Brown-Peterson paradigm). This could lead to active displacement of items from the focus according to views (e.g., McElree 2001) that posit such displacement as a mechanism of STM forgetting, or increase the noise according to interference-based explanations (see discussion below in What Happens Neurally During the Delay?). By either account, the problem with retention-interval tasks is that they are questionable ways to prevent rehearsal of the to-be-remembered information, and they introduce new, distracting information that may engage STM. This double-edged sword makes it difficult to tie

retention-interval manipulations directly to decay.

Attempts to address the confounding factors. A potential way out of the rehearsal conundrum is to use stimuli that are not easily converted to verbal codes and that therefore may be difficult to rehearse. For example, Harris (1952) used tones that differed so subtly in pitch that they would be difficult to name by subjects without perfect pitch. On each trial, participants were first presented with a to-be-remembered tone, followed by a retention interval of 0.1 to 25 seconds, and finally a probe tone. The accuracy of deciding whether the initial and probe tones were the same declined with longer retention intervals, consistent with the predictions of decay theory.

Using another technique, McKone (1995, 1998) reduced the probability of rehearsal or other explicit-memory strategies by using an implicit task. Words and nonwords were repeated in a lexical-decision task, with the measure of memory being faster performance on repeated trials than on novel ones (priming). To disentangle the effects of decay and interference, McKone varied the time between repetitions (the decay-related variable) while holding the number of items between repetitions (the interference-related variable) constant, and vice versa. She found that greater time between repetitions reduced priming even after accounting for the effects of intervening items, consistent with decay theory. However, interference and decay effects seemed to interact and to be especially important for nonwords.

Procedures such as those used by Harris (1952) and McKone (1995, 1998) do not have the problems associated with retention-interval tasks. They are, however, potentially vulnerable to the criticism of Keppel & Underwood (1962) regarding interference from prior trials within the task, although McKone's experiments address this issue to some degree. Another potential problem is that these participants' brains and minds are

not inactive during the retention interval (Raichle et al. 2001). There is increasing evidence that the processes ongoing during nominal "resting states" are related to memory, including STM (Hampson et al. 2006). Spontaneous retrieval by participants during the retention interval could interfere with memory for the experimental items. So, although experiments that reduce the influence of rehearsal provide some of the best evidence of decay, they are not definitive.

What happens neurally during the delay?

Neural findings of delay-period activity have also been used to support the idea of decay. For example, at the single-cell level, Fuster (1995) found that in monkeys performing a delayed-response task, delay-period activity in inferotemporal cortex steadily declined over 18 seconds (see also Pasternak & Greenlee 2005). At a molar level, human neuroimaging studies often show delay-period activity in prefrontal and posterior regions, and this activity is often thought to support maintenance or storage (see review by Smith & Jonides 1999). As reviewed above, it is likely that the posterior regions support storage and that frontal regions support processes related to interference-resolution, control, attention, response preparation, motivation, and reward.

Consistent with the suggestive primate data, Jha & McCarthy (2000) found a general decline in activation in posterior regions over a delay period, which suggests some neural evidence for decay. However, this decline in activation was not obviously related to performance, which suggests two (not mutually exclusive) possibilities: (a) the decline in activation was not representative of decay, so it did not correlate with performance, or (b) these regions might not have been storage regions (but see Todd & Marois 2004 and Xu & Chun 2006 for evidence more supportive of load sensitivity in posterior regions).

The idea that neural activity decays also faces a serious challenge in the classic results of Malmö (1942), who found that a monkey

with frontal lesions was able to perform a delayed response task extremely well (97% correct) if visual stimulation and motor movement (and therefore associated interference) were restricted during a 10-second delay. By contrast, in unrestricted conditions, performance was as low as 25% correct (see also Postle & D'Esposito 1999). In summary, evidence for time-based declines in neural activity that would naturally be thought to be part of a decay process is at best mixed.

Is there a mechanism for decay? Although there are data supporting the existence of decay, much of these data are subject to alternative, interference-based explanations. However, as Crowder (1976) noted, "Good ideas die hard." At least a few key empirical results (Harris 1952; McKone 1995, 1998) do seem to implicate some kind of time-dependent decay. If one assumes that decay happens, how might it occur?

One possibility—perhaps most compatible with results like those of Malmö (1942)—is that what changes over time is not the integrity of the representation itself, but the likelihood that attention will be attracted away from it. As more time passes, the likelihood increases that attention will be attracted away from the target and toward external stimuli or other memories, and it will be more difficult to return to the target representation. This explanation seems compatible with the focus-of-attention views of STM that we have reviewed. By this explanation, capacity limits are a function of attention limits rather than a special property of STM per se.

Another explanation, perhaps complementary to the first, relies on stochastic variability in the neuronal firing patterns that make up the target representation. The temporal synchronization of neuronal activity is an important part of the representation (e.g., Deiber et al. 2007, Jensen 2006, Lisman & Idiart 1995). As time passes, variability in the firing rates of individual neurons may cause them to fall increasingly out of synchrony unless they are reset (e.g., by rehearsal). As

the neurons fall out of synchrony, by this hypothesis, the firing pattern that makes up the representation becomes increasingly difficult to discriminate from surrounding noise [see Lustig et al. (2005) for an example that integrates neural findings with computational (Frank et al. 2001) and behaviorally based (Brown et al. 2000) models of STM].

Interference Theories: Comprehensive but Complex

Interference effects play several roles in memory theory: First, they are the dominant explanation of forgetting. Second, some have suggested that STM capacity and its variation among individuals is largely determined by the ability to overcome interference (e.g., Hasher & Zacks 1988, Unsworth & Engle 2007). Finally, differential interference effects in STM and LTM have been used to justify the idea that they are separate systems, and common interference effects have been used to justify the idea that they are a unitary system.

Interference theory has the opposite problem of decay: It is comprehensive but complex (Crowder 1976). The basic principles are straightforward. Items in memory compete, with the amount of interference determined by the similarity, number, and strength of the competitors. The complexity stems from the fact that interference may occur at multiple stages (encoding, retrieval, and possibly storage) and at multiple levels (the representation itself or its association with a cue or a response). Interference from the past (proactive interference; PI) may affect both the encoding and the retrieval of new items, and it often increases over time. By contrast, interference from new items onto older memories (retroactive interference; RI) frequently decreases over time and may not be as reliant on similarity (see discussion by Wixted 2004).

Below, we review some of the major findings with regard to interference in STM, including a discussion of its weaknesses in

explaining short-term forgetting. We then present a conceptual model of STM that attempts to address these weaknesses and the questions regarding structure, process, and forgetting raised throughout this review.

Interference Effects in Short-Term Memory

Selection-based interference effects. The Brown-Peterson task, originally conceived to test decay theory, became a workhorse for testing similarity-based interference as well. In the “release-from-PI” version (Wickens 1970), short lists of categorized words are used as memoranda. Participants learn one three-item list on each trial, perform some other task during the retention interval, and then attempt to recall the list. For the first three trials, all lists consist of words from the same category (e.g., flowers). The typical PI effects occur: Recall declines over subsequent trials. The critical manipulation occurs at the final list. If it is from a different category (e.g., sports), recall is much higher than if it is from the same category as preceding trials. In some cases, performance on this set-shift or release-from-PI trial is nearly as high as on the very first trial.

The release-from-PI effect was originally interpreted as an encoding effect. Even very subtle shifts (e.g., from “flowers” to “wild-flowers”) produce the effect if participants are warned about the shift before the words are presented (see Wickens 1970 for an explanation). However, Gardiner et al. (1972) showed that release also occurs if the shift-cue is presented only at the time of the retrieval test—i.e., after the list has been encoded. They suggested that cues at retrieval could reduce PI by differentiating items from the most recent list, thus aiding their selection.

Selection processes remain an important topic in interference research. Functional neuroimaging studies consistently identify a region in left inferior frontal gyrus (LIFG) as active during interference resolution, at least for verbal materials (see

a review by Jonides & Nee 2006). This region appears to be generally important for selection among competing alternatives, e.g., in semantic memory as well as in STM (Thompson-Schill et al. 1997). In STM, LIFG is most prominent during the test phase of interference trials, and its activation during this phase often correlates with behavioral measures of interference resolution (D'Esposito et al. 1999, Jonides et al. 1998, Reuter-Lorenz et al. 2000, Thompson-Schill et al. 2002). These findings attest to the importance of processes for resolving retrieval interference. The commonality of the neural substrate for interference resolution across short-term and long-term tasks provides yet further support for the hypothesis of shared retrieval processes for the two types of memory.

Interference effects occur at multiple levels, and it is important to distinguish between interference at the level of representations and interference at the level of responses. The LIFG effects described above appear to be familiarity based and to occur at the level of representations. Items on a current trial must be distinguished and selected from among items on previous trials that are familiar because of prior exposure but are currently incorrect. A separate contribution occurs at the level of responses: An item associated with a positive response on a prior trial may now be associated with a negative response, or vice versa. This response-based conflict can be separated from the familiarity-based conflict, and its resolution appears to rely more on the anterior cingulate (Nelson et al. 2003).

Other mechanisms for interference effects? Despite the early work of Keppel & Underwood (1962), most studies examining encoding in STM have focused on RI: how new information disrupts previous memories. Early theorists described this disruption in terms of displacement of entire items from STM, perhaps by disrupting consolidation (e.g., Waugh & Norman 1965). However, rapid serial visual presentation studies sug-

gest that this type of consolidation is complete within a very short time—approximately 500 milliseconds, and in some situations as short as 50 milliseconds (Vogel et al. 2006).

What about interference effects beyond this time window? As reviewed above, most current focus-based models implicitly assume something like whole-item displacement is at work, but these models may need to be elaborated to account for retroactive similarity-based interference, such as the phonological interference effects reviewed by Nairne (2002). The models of Nairne (2002) and Oberauer (2006) suggest a direction for such an elaboration. Rather than a competition at the item level for a single-focus resource, these models posit a lower-level similarity-based competition for “feature units.” By this idea, items in STM are represented as bundles of features (e.g., color, shape, spatial location, temporal location). Representations of these features in turn are distributed over multiple units. The more two items overlap, the more they compete for these feature units, resulting in greater interference. This proposed mechanism fits well with the idea that working memory reflects the heightened activation of representations that are distributed throughout sensory, semantic, and motor cortex (Postle 2006), and that similarity-based interference constrains the capacity due to focusing (see above; Awh et al. 2007). Hence, rather than whole-item displacement, specific feature competition may underlie the majority of encoding-stage RI.

Interference-based decay? Above, we proposed a mechanism for decay based on the idea that stochastic variability causes the neurons making up a representation to fall out of synchrony (become less coherent in their firing patterns). Using the terminology of Nairne (2002) and Oberauer (2006), the feature units become less tightly bound. Importantly, feature units that are not part of a representation also show some random activity due to their own stochastic variability, creating a noise

distribution. Over time, there is an increasing likelihood that the feature units making up the to-be-remembered item's representation will overlap with those of the noise distribution, making them increasingly difficult to distinguish. This increasing overlap with the noise distribution and loss of feature binding could lead to the smooth forgetting functions often interpreted as evidence for decay.

Such a mechanism for decay has interesting implications. It may explain why PI effects interact with retention interval. Prior trials with similar items would structure the noise distribution so that it is no longer random but rather is biased to share components with the representation of the to-be-remembered item (target). Representations of prior, now-irrelevant items might compete with the current target's representation for control of shared feature units, increasing the likelihood (rate) at which these units fall out of synchrony.

Prior similar items may also dampen the fidelity of the target representation to begin with, weakening their initial binding and thus causing these items to fall out of synchrony more quickly. In addition, poorly learned items might have fewer differentiating feature units, and these units may be less tightly bound and therefore more vulnerable to falling out of synchrony. This could explain why Keppel & Underwood (1962) found that poorly learned items resulted in retention-interval effects even on the first trial. It may also underlie the greater decay effects that McKone (1995, 1998) found for nonwords than for words, if one assumes that nonwords have fewer meaning-based units and connections.

A SUMMARY OF PRINCIPLES AND AN ILLUSTRATION OF SHORT-TERM MEMORY AT WORK

Here we summarize the principles of STM that seem best supported by the behavioral and neural evidence. Building on these principles,

we offer a hypothetical sketch of the processes and neural structures that are engaged by a canonical STM task, the probe recognition task with distracting material.

Principles of Short-Term Memory

We have motivated our review by questions of structure, process, and forgetting. Rather than organize our summary this way, we wish to return here to the title of our review and consider what psychological and neural mechanisms seem best defended by empirical work. In that we have provided details about each of these issues in our main discussion, we summarize them here as bullet points. Taken together, they provide answers to our questions about structure, process, and forgetting.

The mind of short-term memory. Representations in memory are composed of bundles of features for stored information, including features representing the context in which that information was encountered.

- Representations in memory vary in activation, with a dormant state characterizing long-term memories, and varying states of activation due to recent perceptions or retrievals of those representations.
- There is a focus of attention in which a bound collection of information may be held in a state that makes it immediately available for cognitive action. Attention may be focused on only a single chunk of information at a time, where a chunk is defined as a set of items that are bound by a common functional context.
- Items may enter the focus of attention via perceptual encoding or via cue-based retrieval from LTM.
- Items are maintained in the focus via a controlled process of maintenance, with rehearsal being a case of controlled sequential allocation of attentional focus.
- Forgetting occurs when items leave the focus of attention and must compete

with other items to regain the focus (interference), or when the fidelity of the representation declines over time due to stochastic processes (decay).

The brain of short-term memory. Items in the focus of attention are represented by patterns of heightened, synchronized firing of neurons in primary and secondary association cortex.

- The sensorimotor features of items in the focus of attention or those in a heightened state of activation are the same as those activated by perception or action. Information within a representation is associated with the cortical region that houses it (e.g., verbal, spatial, motor). In short, item representations are stored where they are processed.
- Medial temporal structures are important for binding items to their context for both the short- and long-term and for retrieving items whose context is no longer in the focus of attention or not yet fully consolidated in the neocortex.
- The capacity to focus attention is constrained by parietal and frontal mechanisms that modulate processing as well as by increased noise in the neural patterns arising from similarity-based interference or from stochastic variability in firing.
- Frontal structures support controlled processes of retrieval and interference resolution.
- Placing an item into the focus of attention from LTM involves reactivating the representation that is encoded in patterns of neural connection weights.
- Decay arises from the inherent variability of the neural firing of feature bundles that build a representation: The likelihood that the firing of multiple features will fall out of synchrony increases with time due to stochastic variability.

A Sketch of Short-Term Memory at Work

The theoretical principles outlined above summarize our knowledge of the psychological and neural bases of STM, but further insight can be gained by attempting to see how these mechanisms might work together, moment-by-moment, to accomplish the demands of simple tasks. We believe that working through an illustration will not only help to clarify the nature of the proposed mechanisms, but it may also lead to a picture of STM that is more detailed in its bridging of neural process and psychological function.

Toward these ends, we present here a specific implementation of the principles that allows us to give a description of the mechanisms that might be engaged at each point in a simple visual STM task. This exercise leads us to a view of STM that is heavily grounded in concepts of neural activation and plasticity. More specifically, we complement the assumptions about cognitive and brain function above with simple hypotheses about the relative supporting roles of neuronal firing and plasticity (described below). Although somewhat speculative in nature, this description is consistent with the summary principles, and it grounds the approach more completely in a plausible neural model. In particular, it has the virtue of providing an unbroken chain of biological mechanisms that supports the encoding of short-term memories over time.

Figure 1 (see color insert) traces the representation of one item in memory over the course of a few seconds in our hypothetical task. The cognitive events are demarcated at the top of the figure, and the task events at the bottom. In the hypothetical task, the subject must keep track of three visual items (such as novel shapes). The first item is presented for 700 milliseconds, followed by a delay of 2 seconds. The second stimulus then appears, followed by a delay of a few seconds, then the third stimulus, and another delay. Finally, the probe appears, and contact must be made with

the memory for the first item. The assumption is that subjects will engage in a strategy of actively maintaining each item during the delay periods.

Before walking through the timeline in **Figure 1**, let us take a high-level view. At any given time point, a vertical slice through the figure is intended to convey two key aspects of the neural basis of the memory. The first is the extent to which multiple cortical areas contribute to the representation of the item, as indicated by the colored layers corresponding to different cortical areas. The dynamic nature of the relative sizes of the layers captures several of our theoretical assumptions concerning the evolving contribution of those different areas at different functional stages of STM. The second key aspect is the distinction between memory supported by a coherent pattern of active neural firing (captured in solid layers) and memory supported by synaptic plasticity (captured in the hashed layers) (Fuster 2003, Grossberg 2003, Rolls 2000). The simple hypothesis represented here is that perceptual encoding and active-focus maintenance are supported by neuronal firing, and memory of items outside the focus is supported by short-term synaptic plasticity (Zucker & Regehr 2002).³

We now follow the time course of the neural representation of the first item (in the order indicated by the numbers in the figure). (1) The stimulus is presented and rapidly triggers a coherent pattern of activity in posterior perceptual regions, representing both low-level visual features of the item content and its abstract identification in higher-level regions. (2) There is also a rapid onset of the representation of item-context binding (temporal con-

text in our example) supported by the medial-temporal lobes (see section titled "Contesting the Idea of Separate Long-Term and Short-Term Systems") (Ranganath & Blumenfeld 2005). (3) Over the first few hundred milliseconds, this pattern increases in quality, yielding speed-accuracy tradeoffs in perceptual identification. (4) Concurrent with the active firing driven by the stimulus, very short-term synaptic plasticity across cortical areas begins to encode the item's features and its binding to context. Zucker & Regehr (2002) identify at least three distinct plasticity mechanisms that begin to operate on this time scale (tens of milliseconds) and that together are sufficient to produce memories lasting several seconds. (For the use of this mechanism in a prominent neural network model of STM, see Burgess & Hitch 1999, 2005, 2006.) (5) At the offset of the stimulus, the active firing pattern decays very rapidly (consistent with identified mechanisms of rapid decay in short-term potentiation; Zucker & Regehr 2002), but (6) active maintenance, mediated by increased activity in frontal and parietal systems, maintains the firing pattern during the delay period (see sections titled "The Architecture of Unitary-Store Models" and "Maintenance of Items in the Focus") (Pasternak & Greenlee 2005, Ranganath 2006, Ruchkin et al. 2003). This active delay firing includes sustained contribution of MTL to item-context binding (see section titled "Contesting the Idea of Separate Long-Term and Short-Term Systems"). Significant reduction in coherence of the firing pattern may occur as a result of stochastic drift as outlined above (in sections titled "What Happens Neurally During the Delay?" and "Interference-Based Decay?"), possibly leading to a kind of short-term decay during maintenance (see section titled "What Happens Neurally During the Delay?") (Fuster 1995, Pasternak & Greenlee 2005). (7) The active maintenance involves the reuse of posterior perceptual regions in the service of the task demands on STM. This reuse includes even early perceptual areas, but we show

³The alternative to this strong claim is that memory items outside the focus might also be supported by residual active firing. The empirical results reviewed above indicating load-dependent posterior activation might lend support to this alternative if one assumes that the memory load in those experiments was not entirely held in the focus, and that these activations exclusively index firing associated with the memory load itself.

here a drop in the contribution of primary perceptual regions to maintenance in order to indicate a relatively greater effect of top-down control on the later high-level regions (Postle 2006, Ranganath 2006). (8) During this delay period of active maintenance, short-term potentiation continues to lay down a trace of the item and its binding to context via connection weights both within and across cortical regions. The overall efficacy of this memory encoding is the result of the interaction of the possibly decaying active firing pattern with the multiple plasticity mechanisms and their individual facilitation and depression profiles (Zucker & Regehr 2002).

(9) At the end of the delay period and the onset of the second stimulus, the focus rapidly shifts to the new stimulus, and the active firing of the neural pattern of the target stimulus ceases. (10) The memory of the item is now carried completely by the changed synaptic weights, but this change is partially disrupted by the incoming item and its engagement of a similar set of neural activity patterns. Cognitively, this disruption yields similarity-based retroactive interference (see "Other Mechanisms for Interference Effects?") (Nairne 2002). (11) Even in the absence of interference, a variety of biochemical processes give rise to the decay of short-term neural change and therefore the gradual loss of the memory trace over time. This pattern of interference and decay continues during processing of both the second and third stimulus. The probe triggers a rapid memory retrieval of the target item (12), mediated in part by strategic frontal control (see "Neural Mechanisms of Short- and Long-Term Memory Retrieval") (Cabeza et al. 2002, Ranganath et al. 2004). This rapid retrieval corresponds to the re-instantiation of the target item's firing pattern in both posterior perceptual areas (13) and medial-temporal regions, the latter supporting the contextual binding. A plausible neural mechanism for the recovery of this activity pattern at retrieval is the emergent pattern-completion property of attractor networks

(Hopfield 1982). Attractor networks depend on memories encoded in a pattern of connection weights, whose formation and dynamics we have sketched above in terms of short-term synaptic plasticity. Such networks also naturally give rise to the kind of similarity-based proactive interference clearly evident in STM retrieval (see "Selection-Based Interference Effects") (Jonides & Nee 2006, Keppel & Underwood 1962).

We have intentionally left underspecified a precise quantitative interpretation of the *y*-axis in **Figure 1**. Psychologically, it perhaps corresponds to a combination of availability (largely driven by the dichotomous nature of the focus state) and accessibility (driven by a combination of both firing and plasticity). Neurally, it perhaps corresponds to some measure of both firing amplitude and coherence and potential firing amplitude and coherence.

We are clearly a long way from generating something like the plot in **Figure 1** from neuroimaging data on actual tasks—though plots of event-related potentials in STM tasks give us an idea of what these data may look like (Ruchkin et al. 2003). There no doubt is more missing from **Figure 1** than is included (e.g., the role of subcortical structures such as the basal ganglia in the frontal/parietal mediated control, or the reciprocal cortical-thalamic circuits that shape the nature of the neocortical patterns). We nevertheless believe that the time course sketched in **Figure 1** is useful for making clear many of the central properties that characterize the psychological and neural theory of human STM outlined above: (a) STM engages essentially all cortical areas—including medial temporal lobes—and does so from the earliest moments, though it engages these areas differentially at different functional stages. (b) STM reuses the same posterior cortical areas and representations that subserve perception, and active maintenance of these representations depends on these posterior areas receiving input from frontal-parietal circuits. (c) Focused items are

distinguished both functionally and neurally by active firing patterns, and nonfocused memories depend on synaptic potentiation and thereby suffer from decay and retroactive interference. (*d*) Nonfocused memories are reinstated into active firing states via an associative retrieval process subject to proactive interference from similarly encoded patterns.

Postscript: Revisiting Complex Cognition

A major goal of this review has been to bring together psychological theorizing (the mind) and neuroscientific evidence (the brain) of STM. However, any celebration of this union is premature until we address this question: Can our account explain how the mind and brain accomplish the everyday tasks (e.g., completing a tax form) that opened this review? The recognition probe task used in our example and the other procedures discussed throughout the main text are considerably simpler than those everyday tasks. Is it plausible to believe that the system outlined here, particularly in light of its severely limited capacity, could support human cognition in the wild?

It is sobering to note that Broadbent (1993) and Newell (1973, 1990) asked this question nearly two decades ago, and at that time they were considering models of STM with even larger capacities than the one advocated here. Even so, both observed that none of the extant computational models of complex cognitive tasks (e.g., the Newell & Simon 1972 models of problem solving) used contemporary psychological theories of STM. Instead, the complex-cognition models assumed much larger (in some cases, effectively unlimited) working memories. The functional viability of the STM theories of that time was thus never clearly demonstrated. Since then, estimates of STM capacity have only grown smaller, so the question, it would seem, has grown correspondingly more pressing.

Fortunately, cognitive modeling and cognitive theory have also developed over that time, and in ways that would have pleased both Broadbent and Newell. Importantly, many computational cognitive architectures now make assumptions about STM capacity that are congruent with the STM models discussed in this review. The most prominent example is ACT-R, a descendent of the early Newell production-system models. ACT-R continues to serve as the basis of computational models of problem solving (e.g., Anderson & Douglass 2001), sentence processing (Lewis & Vasishth 2005, Lewis et al. 2006), and complex interactive tasks (Anderson et al. 2004). However, the current version of ACT-R has a focus-based structure with an effective capacity limit of four or fewer items (Anderson et al. 2004).

Another important theoretical development is the long-term working memory approach of Ericsson & Kintsch (1995). This approach describes how LTM, using the kind of fast-encoding and cue-based associative retrieval processes assumed here, can support a variety of complex cognitive tasks ranging from discourse comprehension to specialized expert performance. In both the modern approaches to computational architecture and long-term working memory, the power of cognition resides not in capacious short-term buffers but rather in the effective use of an associative LTM. A sharply limited focus of attention does not, after all, seem to pose insurmountable functional problems.

In summary, this review describes the still-developing convergence of computational models of complex cognition, neural network models of simple memory tasks, modern psychological studies of STM, and neural studies of memory in both humans and primates. The points of contact among these different methods of studying STM have multiplied over the past several years. As we have pointed out, significant and exciting challenges in furthering this integration lie ahead.

LITERATURE CITED

- Altmann EM, Gray WD. 2002. Forgetting to remember: the functional relationship of decay and interference. *Psychol. Sci.* 13(1):27–33
- Anderson JR. 1983. Retrieval of information from long-term memory. *Science* 220(4592):25–30
- Anderson JR, Bothell D, Byrne MD, Douglass S, Lebiere C, Qin Y. 2004. An integrated theory of mind. *Psychol. Rev.* 111:1036–60
- Anderson JR, Douglass S. 2001. Tower of Hanoi: evidence for the cost of goal retrieval. *J. Exp. Psychol.: Learn. Mem. Cogn.* 27:1331–46
- Anderson JR, Matessa M. 1997. A production system theory of serial memory. *Psychol. Rev.* 104(4):728–48
- Anderson JR, Schooler LJ. 1991. Reflections of the environment in memory. *Psychol. Sci.* 2(6):396–408
- Atkinson RC, Shiffrin RM. 1971. The control of short-term memory. *Sci. Am.* 224:82–90
- Awh E, Barton B, Vogel EK. 2007. Visual working memory represents a fixed number of items regardless of complexity. *Psychol. Sci.* 18(7):622–28
- Awh E, Jonides J. 2001. Overlapping mechanisms of attention and spatial working memory. *Trends Cogn. Sci.* 5(3):119–26
- Awh E, Jonides J, Reuter-Lorenz PA. 1998. Rehearsal in spatial working memory. *J. Exp. Psychol.: Hum. Percept. Perform.* 24:780–90
- Awh E, Jonides J, Smith EE, Buxton RB, Frank LR, et al. 1999. Rehearsal in spatial working memory: evidence from neuroimaging. *Psychol. Sci.* 10(5):433–37
- Awh E, Jonides J, Smith EE, Schumacher EH, Koeppe RA, Katz S. 1996. Dissociation of storage and rehearsal in verbal working memory: evidence from PET. *Psychol. Sci.* 7:25–31
- Baddeley AD. 1986. *Working Memory*. Oxford: Clarendon
- Baddeley AD. 1992. Working memory. *Science* 225:556–59
- Baddeley AD. 2000. The episodic buffer: a new component of working memory? *Trends Cogn. Sci.* 4(11):417–23
- Baddeley AD. 2003. Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4(10):829–39
- Baddeley AD, Hitch G. 1974. Working memory. In *Recent Advances in Learning and Motivation*, Vol. 8, ed. GA Bower, pp. 47–90. New York: Academic
- Baddeley AD, Thomson N, Buchanan M. 1975. Word length and structure of short-term memory. *J. Verbal Learn. Verbal Behav.* 14(6):575–89
- Baddeley AD, Vallar G, Wilson BA. 1987. Sentence comprehension and phonological working memory: some neuropsychological evidence. In *Attention and Performance XII: The Psychology of Reading*, ed. M Coltheart, pp. 509–29. London: Erlbaum
- Baddeley AD, Warrington EK. 1970. Amnesia and the distinction between long- and short-term memory. *J. Verbal Learn. Verbal Behav.* 9:176–89
- Baddeley AD, Wilson BA. 2002. Prose recall and amnesia: implications for the structure of working memory. *Neuropsychologia* 40:1737–43
- Badre D, Wagner AD. 2005. Frontal lobe mechanisms that resolve proactive interference. *Cereb. Cortex* 15:2003–12
- Braver TS, Barch DM, Kelley WM, Buckner RL, Cohen NJ, et al. 2001. Direct comparison of prefrontal cortex regions engaged by working and long-term memory tasks. *Neuroimage* 14:48–59
- Broadbent D. 1993. Comparison with human experiments. In *The Simulation of Human Intelligence*, ed. D Broadbent, pp. 198–217. Oxford: Blackwell Sci.
- Brooks LR. 1968. Spatial and verbal components of the act of recall. *Can. J. Psychol.* 22:349–68

- Brown GDA, Preece T, Hulme C. 2000. Oscillator-based memory for serial order. *Psychol. Rev.* 107(1):127–81
- Buckner RL, Koutstaal W, Schacter DL, Wagner AD, Rosen BR. 1998. Functional-anatomic study of episodic retrieval using fMRI: I. Retrieval effort versus retrieval success. *NeuroImage* 7(3):151–62
- Buffalo EA, Reber PJ, Squire LR. 1998. The human perirhinal cortex and recognition memory. *Hippocampus* 8:330–39
- Burgess N, Hitch GJ. 1999. Memory for serial order: a network model of the phonological loop and its timing. *Psychol. Rev.* 106(3):551–81
- Burgess N, Hitch GJ. 2005. Computational models of working memory: putting long-term memory into context. *Trends Cogn. Sci.* 9:535–41
- Burgess N, Hitch GJ. 2006. A revised model of short-term memory and long-term learning of verbal sequences. *J. Mem. Lang.* 55:627–52
- Cabeza R, Dolcos F, Graham R, Nyberg L. 2002. Similarities and differences in the neural correlates of episodic memory retrieval and working memory. *Neuroimage* 16:317–30
- Cabeza R, Nyberg L. 2000. Imaging cognition II: an empirical review of 275 PET and fMRI studies. *J. Cogn. Neurosci.* 9:254–65
- Cave CB, Squire LR. 1992. Intact verbal and nonverbal short-term memory following damage to the human hippocampus. *Hippocampus* 2:151–63
- Cowan N. 1988. Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychol. Bull.* 104:163–91
- Cowan N. 1995. *Attention and Memory: An Integrated Framework*. New York: Oxford Univ. Press
- Cowan N. 2000. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24:87–185
- Crowder R. 1976. *Principles of Learning and Memory*. Hillsdale, NJ: Erlbaum
- Damasio AR. 1989. Time-locked multiregional retroactivation: a system-level proposal for the neuronal substrates of recall and recognition. *Cognition* 33:25–62
- Darwin CJ, Turvey MT, Crowder RG. 1972. Auditory analogue of Sperling partial report procedure—evidence for brief auditory storage. *Cogn. Psychol.* 3(2):255–67
- Deiber MP, Missonnier P, Bertrand O, Gold G, Fazio-Costa L, et al. 2007. Distinction between perceptual and attentional processing in working memory tasks: a study of phase-locked and induced oscillatory brain dynamics. *J. Cogn. Neurosci.* 19(1):158–72
- den Heyer K, Barrett B. 1971. Selective loss of visual and verbal information in STM by means of visual and verbal interpolated tasks. *Psychon. Sci.* 25:100–2
- D'Esposito M, Postle BR. 1999. The dependence of span and delayed-response performance on prefrontal cortex. *Neuropsychologia* 37(11):1303–15
- D'Esposito M, Postle BR. 2000. Neural correlates of processes contributing to working memory function: evidence from neuropsychological and pharmacological studies. In *Control of Cognitive Processes*, ed. S Monsell, J Driver, pp. 580–602. Cambridge, MA: MIT Press
- D'Esposito M, Postle BR, Jonides J, Smith EE, Lease J. 1999. The neural substrate and temporal dynamics of interference effects in working memory as revealed by event-related fMRI. *Proc. Natl. Acad. Sci. USA* 96:7514–19
- Eng HY, Chen DY, Jiang YH. 2005. Visual working memory for simple and complex visual stimuli. *Psychon. Bull. Rev.* 12:1127–33
- Ericsson KA, Kintsch W. 1995. Long-term working memory. *Psychol. Rev.* 102:211–45
- Fletcher PC, Henson RNA. 2001. Frontal lobes and human memory—insights from functional neuroimaging. *Brain* 124:849–81

- Frank MJ, Loughry B, O'Reilly RC. 2001. Interactions between the frontal cortex and basal ganglia in working memory: a computational model. *Cogn. Affect. Behav. Neurosci.* 1:137–60
- Funahashi S, Bruce CJ, Goldman-Rakic PS. 1989. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* 61:331–49
- Fuster JK. 2003. Thoughts from the long-term memory chair. *Behav. Brain Sci.* 26:734–35
- Fuster JM. 1973. Unit activity in prefrontal cortex during delayed response performance: neuronal correlates of transient memory. *J. Neurophysiol.* 36:61–78
- Fuster JM. 1995. *Memory in the Cerebral Cortex*. Cambridge, MA: MIT Press
- Gabrieli JDE, Brewer JB, Desmond JE, Glover GH. 1997. Separate neural bases of two fundamental memory processes in the human medial temporal lobe. *Science* 276:264–66
- Garavan H. 1998. Serial attention within working memory. *Mem. Cogn.* 26:263–76
- Gardiner JM, Craik FIM, Birtwist J. 1972. Retrieval cues and release from proactive inhibition. *J. Verbal Learn. Verbal Behav.* 11(6):778–83
- Gillund G, Shiffrin RM. 1984. A retrieval model for both recognition and recall. *Psychol. Rev.* 91(1):1–67
- Goldman-Rakic PS. 1987. Circuitry of primate pre-frontal cortex and regulation of behavior by representational memory. In *Handbook of Physiology: The Nervous System*, Vol. 5, ed. F. Plum, pp. 373–417. Bethesda, MD: Am. Physiol. Soc.
- Grossberg S. 2003. From working memory to long-term memory and back: linked but distinct. *Behav. Brain Sci.* 26:737–38
- Hampson M, Driesen NR, Skudlarski P, Gore JC, Constable RT. 2006. Brain connectivity related to working memory performance. *J. Neurosci.* 26(51):13338–43
- Hanley JR, Young AW, Pearson NA. 1991. Impairment of the visuo-spatial sketch pad. *Q. J. Exp. Psychol. Hum. Exp. Psychol.* 43:101–25
- Hannula DE, Tranel D, Cohen NJ. 2006. The long and the short of it: relational memory impairments in amnesia, even at short lags. *J. Neurosci.* 26(32):8352–59
- Harris JD. 1952. The decline of pitch discrimination with time. *J. Exp. Psychol.* 43(2):96–99
- Hasher L, Zacks RT. 1988. Working memory, comprehension, and aging: a review and a new view. In *The Psychology of Learning and Motivation*, Vol. 22, ed. GH Bower, pp. 193–225. New York: Academic
- Hebb DO. 1949. *The Organization of Behavior*. New York: Wiley
- Hockley WE. 1984. Analysis of response-time distributions in the study of cognitive-processes. *J. Exp. Psychol.: Learn. Mem. Cogn.* 10(4):598–615
- Holdstock JS, Shaw C, Aggleton JP. 1995. The performance of amnesic subjects on tests of delayed matching-to-sample and delayed matching-to-position. *Neuropsychologia* 33:1583–96
- Hopfield JJ. 1982. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA* 79(8):2554–58
- Jacobsen CF. 1936. The functions of the frontal association areas in monkeys. *Comp. Psychol. Monogr.* 13:1–60
- James W. 1890. *Principles of Psychology*. New York: Henry Holt
- Jensen O. 2006. Maintenance of multiple working memory items by temporal segmentation. *Neuroscience* 139:237–49
- Jha AP, McCarthy G. 2000. The influence of memory load upon delay-interval activity in a working-memory task: an event-related functional MRI study. *J. Cogn. Neurosci.* 12:90–105
- Johnson MK, Raye CL, Mitchell KJ, Greene EJ, Cunningham WA, Sanislow CA. 2005. Using fMRI to investigate a component process of reflection: prefrontal correlates of refreshing a just-activated representation. *Cogn. Affect. Behav. Neurosci.* 5:339–61

- Jonides J, Lacey SC, Nee DE. 2005. Processes of working memory in mind and brain. *Curr. Dir. Psychol. Sci.* 14:2–5
- Jonides J, Nee DE. 2006. Brain mechanisms of proactive interference in working memory. *Neuroscience* 139:181–93
- Jonides J, Smith EE, Koeppe RA, Awh E, Minoshima S, Mintun MA. 1993. Spatial working memory in humans as revealed by PET. *Nature* 363:623–25
- Jonides J, Smith EE, Marshuetz C, Koeppe RA, Reuter-Lorenz PA. 1998. Inhibition in verbal working memory revealed by brain activation. *Proc. Natl. Acad. Sci. USA* 95:8410–13
- Keppel G, Underwood BJ. 1962. Proactive-inhibition in short-term retention of single items. *J. Verbal Learn. Verbal Behav.* 1:153–61
- Lange EB, Oberauer K. 2005. Overwriting of phonemic features in serial recall. *Memory* 13:333–39
- Lewandowsky S, Duncan M, Brown GDA. 2004. Time does not cause forgetting in short-term serial recall. *Psychon. Bull. Rev.* 11:771–90
- Lewis RL, Vasishth S. 2005. An activation-based theory of sentence processing as skilled memory retrieval. *Cogn. Sci.* 29:375–419
- Lewis RL, Vasishth S, Van Dyke J. 2006. Computational principles of working memory in sentence comprehension. *Trends Cogn. Sci.* 10:447–54
- Lisman JE, Idiart MAP. 1995. Storage of 7+/-2 short-term memories in oscillatory subcycles. *Science* 267:1512–15
- Luck SJ, Vogel EK. 1997. The capacity of visual working memory for features and conjunctions. *Nature* 390:279–81
- Lustig C, Matell MS, Meck WH. 2005. Not “just” a coincidence: frontal-striatal interactions in working memory and interval timing. *Memory* 13:441–48
- Malmo RB. 1942. Interference factors in delayed response in monkeys after removal of frontal lobes. *J. Neurophysiol.* 5:295–308
- Manoach DS, Greve DN, Lindgren KA, Dale AM. 2003. Identifying regional activity associated with temporally separated components of working memory using event-related functional MRI. *NeuroImage* 20(3):1670–84
- Martin RC. 1993. Short-term memory and sentence processing: evidence from neuropsychology. *Mem. Cogn.* 21:176–83
- McClelland JL, McNaughton BL, O'Reilly RC. 1995. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* 102:419–57
- McElree B. 1998. Attended and nonattended states in working memory: accessing categorized structures. *J. Mem. Lang.* 38:225–52
- McElree B. 2001. Working memory and focal attention. *J. Exp. Psychol.: Learn. Mem. Cogn.* 27:817–35
- McElree B. 2006. Accessing recent events. *Psychol. Learn. Motiv.* 46:155–200
- McElree B, Doshier BA. 1989. Serial position and set size in short-term memory: time course of recognition. *J. Exp. Psychol.: Gen.* 118:346–73
- McGeoch J. 1932. Forgetting and the law of disuse. *Psychol. Rev.* 39:352–70
- McKone E. 1995. Short-term implicit memory for words and non-words. *J. Exp. Psychol.: Learn. Mem. Cogn.* 21(5):1108–26
- McKone E. 1998. The decay of short-term implicit memory: unpacking lag. *Mem. Cogn.* 26(6):1173–86
- Meyer DE, Kieras DE. 1997. A computational theory of executive cognitive processes and multiple-task performance: 1. Basic mechanisms. *Psychol. Rev.* 104(1):3–65

- Miller GA. 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63:81–97
- Milner PM. 2001. Magical attention. *Behav. Brain Sci.* 24(1):131
- Miyashita Y, Chang HS. 1968. Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature* 331:68–70
- Monsell S. 1978. Recency, immediate recognition memory, and reaction-time. *Cogn. Psychol.* 10(4):465–501
- Murdock BB. 1982. A theory for the storage and retrieval of item and associative information. *Psychol. Rev.* 89(6):609–26
- Nairne JS. 2002. Remembering over the short-term: the case against the standard model. *Annu. Rev. Psychol.* 53:53–81
- Neath I, Nairne JS. 1995. Word-length effects in immediate memory: overwriting trace decay theory. *Psychon. Bull. Rev.* 2:429–41
- Nelson JK, Reuter-Lorenz PA, Sylvester CYC, Jonides J, Smith EE. 2003. Dissociable neural mechanisms underlying response-based and familiarity-based conflict in working memory. *Proc. Natl. Acad. Sci. USA* 100:11171–75
- Newell A. 1973. You can't play 20 questions with nature and win: projective comments on the papers of this symposium. In *Visual Information Processing*, ed. WG Chase, pp. 283–310. New York: Academic
- Newell A. 1990. *Unified Theories of Cognition*. Cambridge, MA: Harvard Univ. Press
- Newell A, Simon H. 1972. *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall
- Nichols EA, Kao Y-C, Verfaellie M, Gabrieli JDE. 2006. Working memory and long-term memory for faces: evidence from fMRI and global amnesia for involvement of the medial temporal lobes. *Hippocampus* 16:604–16
- Oberauer K. 2002. Access to information in working memory: exploring the focus of attention. *J. Exp. Psychol.: Learn. Mem. Cogn.* 28:411–21
- Oberauer K. 2006. Is the focus of attention in working memory expanded through practice? *J. Exp. Psychol.: Learn. Mem. Cogn.* 32:197–214
- Oberauer K, Kliegl R. 2006. A formal model of capacity limits in working memory. *J. Mem. Lang.* 55:601–26
- Olson IR, Moore KS, Stark M, Chatterjee A. 2006a. Visual working memory is impaired when the medial temporal lobe is damaged. *J. Cogn. Neurosci.* 18:1087–97
- Olson IR, Page K, Moore KS, Chatterjee A, Verfaellie M. 2006b. Working memory for conjunctions relies on the medial temporal lobe. *J. Neurosci.* 26:4596–601
- Owen AM, Sahakian BJ, Semple J, Polkey CE, Robbins TW. 1995. Visuo-spatial short-term recognition memory and learning after temporal lobe excisions, frontal lobe excisions or amygdala-hippocampectomy in man. *Neuropsychologia* 33:1–24
- Pashler H. 1988. Familiarity and visual change detection. *Percept. Psychophys.* 44:369–78
- Pasternak T, Greenlee MW. 2005. Working memory in primate sensory systems. *Nat. Rev. Neurosci.* 6:97–107
- Polk TA, Simen P, Lewis RL, Freedman E. 2002. A computational approach to control in complex cognition. *Cogn. Brain Res.* 15(1):71–83
- Postle BR. 2006. Working memory as an emergent property of the mind and brain. *Neuroscience* 139:23–38
- Postle BR, D'Esposito M. 1999. "What"—then—"where" in visual working memory: an event-related, fMRI study. *J. Cogn. Neurosci.* 11(6):585–97
- Postman L. 1961. Extra-experimental interference and retention of words. *J. Exp. Psychol.* 61(2):97–110

- Prabhakaran V, Narayanan ZZ, Gabrieli JDE. 2000. Integration of diverse information in working memory within the frontal lobe. *Nat. Neurosci.* 3:85–90
- Pylyshyn ZW. 1994. Some primitive mechanisms of spatial attention. *Cognition* 50:363–84
- Pylyshyn ZW, Burkell J, Fisher B, Sears C, Schmidt W, Trick L. 1994. Multiple parallel access in visual-attention. *Can. J. Exp. Psychol. Rev. Can. Psychol. Exp.* 48(2):260–83
- Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL. 2001. A default mode of brain function. *Proc. Natl. Acad. Sci. USA* 98:676–82
- Ranganath C. 2006. Working memory for visual objects: complementary roles of inferior temporal, medial temporal, and prefrontal cortex. *Neuroscience* 139:277–89
- Ranganath C, Blumenfeld RS. 2005. Doubts about double dissociations between short- and long-term memory. *Trends Cogn. Sci.* 9:374–80
- Ranganath C, DeGutis J, D'Esposito M. 2004. Category-specific modulation of inferior temporal activity during working memory encoding and maintenance. *Cogn. Brain Res.* 20:37–45
- Ranganath C, D'Esposito M. 2001. Medial temporal lobe activity associated with active maintenance of novel information. *Neuron* 31:865–73
- Ranganath C, D'Esposito M. 2005. Directing the mind's eye: prefrontal, inferior and medial temporal mechanisms for visual working memory. *Curr. Opin. Neurobiol.* 15:175–82
- Ranganath C, Johnson MK, D'Esposito M. 2003. Prefrontal activity associated with working memory and episodic long-term memory. *Neuropsychologia* 41(3):378–89
- Renart A, Parga N, Rolls ET. 1999. Backward projections in the cerebral cortex: implications for memory storage. *Neural Comput.* 11(6):1349–88
- Repp G, Baddeley AD. 2006. The multi-component model of working memory: explorations in experimental cognitive psychology. *Neuroscience* 139:5–21
- Reuter-Lorenz PA, Jonides J. 2007. The executive is central to working memory: insights from age, performance and task variations. In *Variations in Working Memory*, ed. AR Conway, C Jarrold, MJ Kane, A Miyake, JN Towse, pp. 250–70. London/New York: Oxford Univ. Press
- Reuter-Lorenz PA, Jonides J, Smith EE, Hartley A, Miller A, et al. 2000. Age differences in the frontal lateralization of verbal and spatial working memory revealed by PET. *J. Cogn. Neurosci.* 12:174–87
- Roediger HL, Knight JL, Kantowitz BH. 1977. Inferring decay in short-term-memory—the issue of capacity. *Mem. Cogn.* 5(2):167–76
- Rolls ET. 2000. Memory systems in the brain. *Annu. Rev. Psychol.* 51:599–630
- Rougier NP, Noelle DC, Braver TS, Cohen JD, O'Reilly RC. 2005. Prefrontal cortex and flexible cognitive control: rules without symbols. *Proc. Natl. Acad. Sci. USA* 102(20):7338–43
- Ruchkin DS, Grafman J, Cameron K, Berndt RS. 2003. Working memory retention systems: a state of activated long-term memory. *Behav. Brain Sci.* 26:709–77
- Sakai K. 2003. Reactivation of memory: role of medial temporal lobe and prefrontal cortex. *Rev. Neurosci.* 14(3):241–52
- Schubert T, Frensch PA. 2001. How unitary is the capacity-limited attentional focus? *Behav. Brain Sci.* 24(1):146
- Scoville WB, Milner B. 1957. Loss of recent memory after bilateral hippocampal lesions. *J. Neurol. Neurosurg. Psychiatry* 20:11–21
- Shallice T, Warrington EK. 1970. Independent functioning of verbal memory stores: a neuropsychological study. *Q. J. Exp. Psychol.* 22:261–73
- Smith EE, Jonides J. 1997. Working memory: a view from neuroimaging. *Cogn. Psychol.* 33:5–42
- Smith EE, Jonides J. 1999. Neuroscience—storage and executive processes in the frontal lobes. *Science* 283:1657–61

- Smith EE, Jonides J, Koeppe RA, Awh E, Schumacher EH, Minoshima S. 1995. Spatial vs object working-memory: PET investigations. *J. Cogn. Neurosci.* 7:337–56
- Sperling G. 1960. The information available in brief visual presentations. *Psychol. Monogr.* 74:Whole No. 498
- Squire L. 1992. Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychol. Rev.* 99:195–231
- Sternberg S. 1966. High-speed scanning in human memory. *Science* 153:652–54
- Talimi D, Grady CL, Goshen-Gottstein Y, Moscovitch M. 2005. Neuroimaging the serial position curve. *Psychol. Sci.* 16:716–23
- Thompson-Schill SL, D'Esposito M, Aguirre GK, Farah MJ. 1997. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proc. Natl. Acad. Sci. USA* 94:14792–97
- Thompson-Schill SL, Jonides J, Marshuetz C, Smith EE, D'Esposito M, et al. 2002. Effects of frontal lobe damage on interference effects in working memory. *J. Cogn. Affect. Behav. Neurosci.* 2:109–20
- Todd JJ, Marois R. 2004. Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* 428(6984):751–54
- Todd JJ, Marois R. 2005. Posterior parietal cortex activity predicts individual differences in visual short-term memory capacity. *Cogn. Affect. Behav. Neurosci.* 5:144–55
- Trick LM, Pylyshyn ZW. 1993. What enumeration studies can show us about spatial attention—evidence for limited capacity preattentive processing. *J. Exp. Psychol.: Hum. Percept. Perform.* 19(2):331–51
- Ungerleider LG, Haxby JV. 1994. “What” and “where” in the human brain. *Curr. Opin. Neurobiol.* 4:157–65
- Unsworth N, Engle RW. 2007. The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychol. Rev.* 114:104–32
- Vallar G, Baddeley AD. 1984. Fractionation of working memory: neuropsychological evidence for a phonological short-term store. *J. Verbal Learn. Verbal Behav.* 23:151–61
- Vallar G, Papagno C. 2002. Neuropsychological impairments of verbal short-term memory. In *The Handbook of Memory Disorders*, ed. AD Baddeley, MD Kopelman, BA Wilson, pp. 249–70. Chichester, UK: Wiley. 2nd ed.
- Verhaeghen P, Basak C. 2007. Aging and switching of the focus of attention in working memory: results from a modified N-Back task. *Q. J. Exp. Psychol. A*: In press
- Verhaeghen P, Cerella J, Basak C. 2004. A working memory workout: how to expand the focus of serial attention from one to four items in 10 hours or less. *J. Exp. Psychol.: Learn. Mem. Cogn.* 30:1322–37
- Vogel EK, Machizawa MG. 2004. Neural activity predicts individual differences in visual working memory capacity. *Nature* 426:748–51
- Vogel EK, Woodman GF, Luck SJ. 2006. The time course of consolidation in visual working memory. *J. Exp. Psychol.: Hum. Percept. Perform.* 32:1436–51
- Wager TD, Smith EE. 2003. Neuroimaging studies of working memory: a meta-analysis. *Neuroimage* 3:255–74
- Warrington EK, Shallice T. 1969. The selective impairment of auditory verbal short-term memory. *Brain* 92:885–96
- Waugh NC, Norman DA. 1965. Primary memory. *Psychol. Rev.* 72:89–104
- Wheeler ME, Peterson SE, Buckner RL. 2000. Memory's echo: vivid remembering reactivates sensory-specific cortex. *Proc. Natl. Acad. Sci. USA* 97(20):11125–29

- Wheeler ME, Shulman GL, Buckner RL, Miezin FM, Velanova K, Petersen SE. 2006. Evidence for separate perceptual reactivation and search processes during remembering. *Cereb. Cortex* 16(7):949–59
- Wickens DD. 1970. Encoding categories of words—empirical approach to meaning. *Psychol. Rev.* 77:1–15
- Wilken P, Ma WJ. 2004. A detection theory account of change detection. *J. Vis.* 4:1120–35
- Wilson FAW, O'Scalaidhe SP, Goldman-Rakic PS. 1993. Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science* 260:1955–58
- Wixted JT. 2004. The psychology and neuroscience of forgetting. *Annu. Rev. Psychol.* 55:235–69
- Woodman GF, Vogel EK, Luck SJ. 2001. Attention is not unitary. *Behav. Brain Sci.* 24(1):153
- Xu YD, Chun MM. 2006. Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature* 440:91–95
- Yantis S, Serences JT. 2003. Cortical mechanisms of space-based and object-based attentional control. *Curr. Opin. Neurobiol.* 13:187–93
- Zhang D, Zhang X, Sun X, Li Z, Wang Z, et al. 2004. Cross-modal temporal order memory for auditory digits and visual locations: an fMRI study. *Hum. Brain Mapp.* 22:280–89
- Zucker RS, Regehr WG. 2002. Short-term synaptic plasticity. *Annu. Rev. Physiol.* 64:355–405

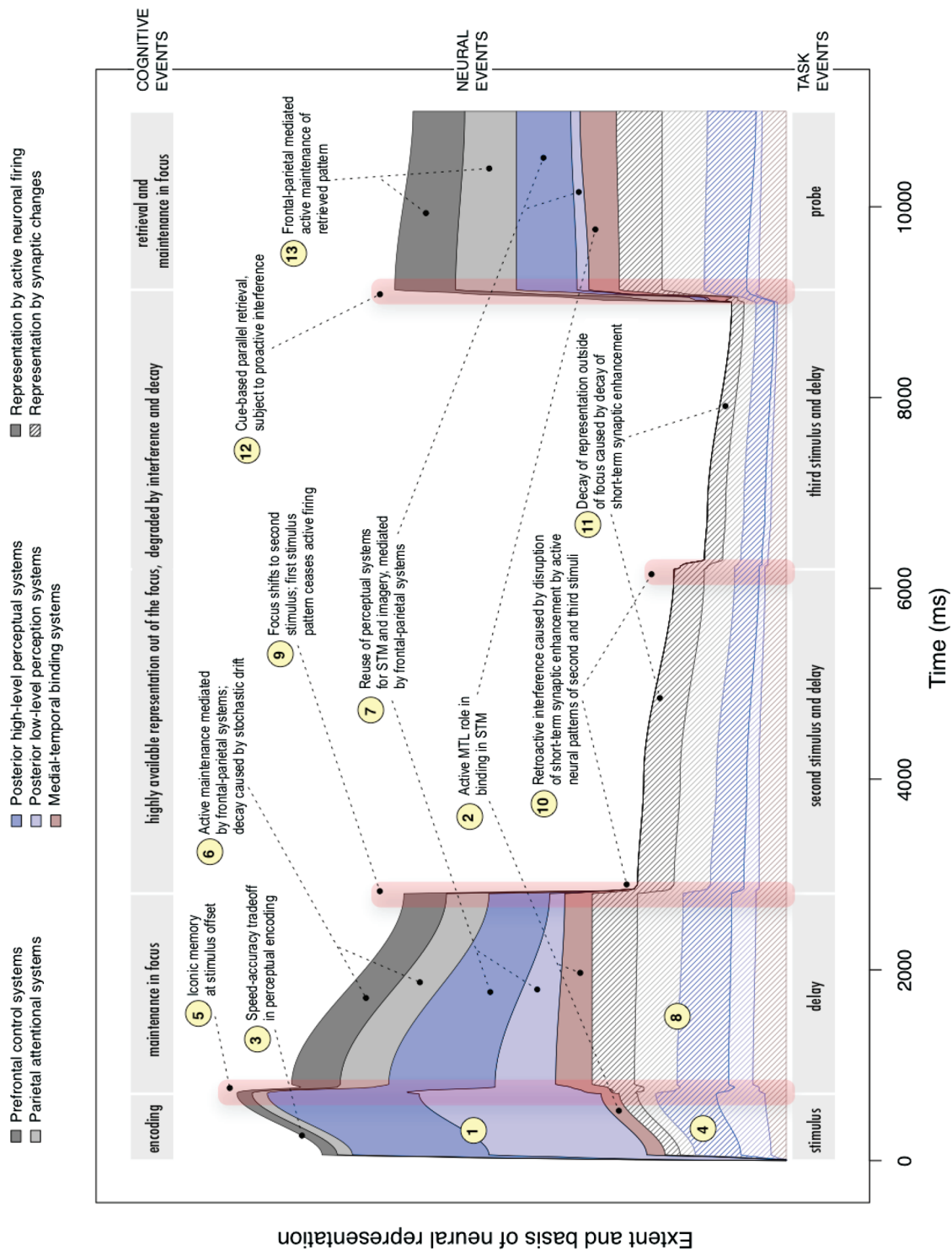


Figure 1

The processing and neural representation of one item in memory over the course of a few seconds in a hypothetical short-term memory task, assuming a simple single-item focus architecture. The cognitive events are demarcated at the top; the task events, at the bottom. The colored layers depict the extent to which different brain areas contribute to the representation of the item over time, at distinct functional stages of short-term memory processing. The colored layers also distinguish two basic types of neural representation: Solid layers depict memory supported by a coherent pattern of active neural firing, and hashed layers depict memory supported by changes in synaptic patterns. The example task requires processing and remembering three visual items; the figure traces the representation of the first item only. In this task, the three items are sequentially presented, and each is followed by a delay period. After the delay following the third item, a probe appears that requires retrieval of the first item. See the text for details corresponding to the numbered steps in the figure.