

Short-Term Memory and Long-Term Memory are Still Different

Dennis Norris

MRC Cognition and Brain Sciences Unit, Cambridge, United Kingdom

A commonly expressed view is that short-term memory (STM) is nothing more than activated long-term memory. If true, this would overturn a central tenet of cognitive psychology—the idea that there are functionally and neurobiologically distinct short- and long-term stores. Here I present an updated case for a separation between short- and long-term stores, focusing on the computational demands placed on any STM system. STM must support memory for previously unencountered information, the storage of multiple tokens of the same type, and variable binding. None of these can be achieved simply by activating long-term memory. For example, even a simple sequence of digits such as “1, 3, 1” where there are 2 tokens of the digit “1” cannot be stored in the correct order simply by activating the representations of the digits “1” and “3” in LTM. I also review recent neuroimaging data that has been presented as evidence that STM is activated LTM and show that these data are exactly what one would expect to see based on a conventional 2-store view.

Keywords: long-term memory, memory, STM, working memory

For more than a century most psychologists have accepted that there are distinct memory systems responsible for long and short-term storage. Originally based entirely on introspection (e.g., James, 1890), the idea that there are separate long- and short-term memory (LTM and STM, respectively) systems subsequently became a core assumption of modern cognitive psychology. From the 1960s most cognitive models of memory have assumed that there are separate stores (Atkinson & Shiffrin, 1968; Baddeley, 1986; Baddeley & Hitch, 1974; Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1992, 2006; Lewandowsky & Farrell, 2000; Page & Norris, 1998a, 1998b, 2009; Waugh & Norman, 1965). This remains the framework guiding almost all cognitive work on verbal STM.

There are dissenting voices. A number of authors have argued that the behavioral data are more parsimoniously accounted for by assuming that there is only a single memory system responsible for both short- and long-term storage (Brown, Neath, & Chater, 2007; Cowan, 1988, 1999; Crowder, 1993; Crowder & Neath, 1991; Engle, Kane, & Tuholski, 1999; Gruneberg, 1970; Jonides et al., 2008; McElree, 2006; Melton, 1963; Nairne, 2002; Poirier & Saint-Aubin, 1996; Surprenant & Neath, 2009). Some have argued that STM is nothing more than activated LTM (Cowan, 1988, 1999; Oberauer, 2002, 2009). This conception of the relation between STM and LTM has become increasingly popular in neu-

roimaging research (Acheson, Hamidi, Binder, & Postle, 2011; Cameron, Haarmann, Grafman, & Ruchkin, 2005; D’Esposito & Postle, 2015; Jonides et al., 2008; LaRocque et al., 2015; Postle, 2006; Ranganath & Blumenfeld, 2005; Ruchkin, Grafman, Cameron, & Berndt, 2003). However, there have been no attempts to present an updated case for a multicomponent view of memory in which there is a clear theoretical and conceptual distinction between LTM and short-term stores. Most of the papers cited above have taken a data-driven approach to the debate, in which involves cataloguing similarities between LTM and STM. Here I take a very different approach and consider the computational requirements of a system that could support retention of information over the short term. These impose strong constraints on the architectural relationship between LTM and STM. The overall conclusion will be that, although LTM has an essential role to play, STM cannot be supported simply by activating LTM. Even simple tasks such as remembering a sequence of digits can only be performed by supplementing LTM with some additional mechanism. This additional mechanism is not part of LTM, but has all of the properties typically ascribed to a separate short-term store. Although I will focus primarily on verbal STM, the underlying logic of the arguments applies equally well to STM in other domains. This article begins with a brief review of the core data supporting a distinction between long- and short-term stores. This is followed by an analysis of cognitive theories that have proposed that STM is activated LTM. The concern here will be to establish exactly what the substantive claims of such theories are. I then review the neuroimaging data that have been presented as support for the activation view. In the second half of the paper I focus on computational considerations that argue strongly against the idea that STM might be nothing more than activated LTM. The central problem for activation-based models is that STM has to be able to store arbitrary configurations of novel information. For example, we can remember novel sequences of words or dots in random positions on a screen. These cannot possibly have preexisting representations in LTM that could be activated. The digit sequence 133646

This article was published Online First May 22, 2017.

This article has been published under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. Copyright for this article is retained by the author(s). Author(s) grant(s) the American Psychological Association the exclusive right to publish the article and identify itself as the original publisher.

Correspondence concerning this article should be addressed to Dennis Norris, MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 2EF, U.K. E-mail: dennis.norris@mrc-cbu.cam.ac.uk

might activate LTM representations of the digits 1, 3, 6 and 4, but some extra mechanism is required to encode the order of the digits and the fact that there are two tokens of the digits 3 and 6. Activating the LTM representations of 1, 3, 6 and 4 is not enough. STM can also store more complex structural representations to encode novel objects and the relation between them. Any viable model of STM must therefore be endowed with additional processes beyond simple activation of LTM. At the very least there must be a mechanism capable of storing multiple tokens, recalling those tokens in the correct order, and creating novel structured representations that cannot yet be in LTM.

I will also ask what it is that we actually store in STM. Does STM contain copies of representations or pointers to representations? If STM contains pointers, do those pointers address LTM, or might those pointers address representations held within the short-term storage system itself? I suggest that, at some level, STM must make use of pointers, but that there are strong computational advantages to having pointers operate within the bounds of a modality-specific STS such as the phonological store component of *Baddeley and Hitch's (1974) Working Memory model*. That is, consideration of how a pointer-based system might work provides further support for the idea that there must be a functional distinction between short- and long-term stores.

Overview

The idea that STM is merely activated LTM is usually pitted against the classic idea that STM consists of a buffer (or buffers) that holds copies of the items in memory. This view of memory is implicit in *Atkinson and Shiffrin's (1966, 1968, 1971) modal model*. Their, 1971 paper begins: "Memory has two components: short-term and long-term" (p. 3). The Atkinson and Shiffrin model is shown in *Figure 1A*. *Figure 1B* shows what is probably the most influential multistore model—the Working Memory model of *Baddeley and Hitch (Baddeley, 2000; Baddeley & Hitch, 1974)*. I will present a very brief summary of the findings that have been taken as evidence for a separation between a long-term store (LTS) and a short-term store (STS). The primary data come from studies of STM patients. These patients typically have grossly impaired verbal STM capacity of perhaps 2–4 items, combined with relatively intact LTM (*Basso, Spinnler, Vallar, & Zanobio, 1982; Saffran & Marin, 1975; Shallice & Vallar, 1990; Shallice & Warrington, 1970; Vallar & Baddeley, 1984; Vallar, Di Betta, & Silveri, 1997; Vallar & Papagno, 1995; Vallar, Papagno, & Baddeley, 1991; Warrington, Logue, & Pratt, 1971; Warrington & Shallice, 1969*). Where it has been examined, these patients generally have few problems with visuospatial STM tasks. The opposite pattern has also been observed. *Hanley, Young, and Pearson (1991)* reported a patient with impaired visuospatial STM but normal verbal STM. One source of evidence that STM patients have intact LTM is that, in general, these patients have little difficulty in learning new associations between familiar stimuli. However, they do have problems in learning new words (*Baddeley, Papagno, & Vallar, 1988; Basso et al., 1982; Bormann, Seyboth, Umarova, & Weiller, 2015; Dittmann & Abel, 2010; Freedman & Martin, 2001; Trojano & Grossi, 1995; Trojano, Stanzione, & Grossi, 1992*). The standard interpretation of this finding is that the complete phonological representation of new words must be held in a short-term store in order to be encoded into LTM. Indeed,

one of the central functions of STM is assumed to be to hold representations that do not yet exist in LTM (*Baddeley, Gathercole, & Papagno, 1998*). Importantly, the constituent parts of these representations (perhaps features, phonemes or syllables) must be stored in the correct order. Not surprisingly, these patients also have problems in learning the order of new digit sequences (*Bormann et al.*). The same logic applies; the sequence must be held in STM so that a representation of the sequence can be presented to LTM.

Conversely, there are patients with medial-temporal lobe damage who have impaired LTM but relatively preserved STM (*Baddeley & Warrington, 1970; Cave & Squire, 1992; Drachman & Arbit, 1966; Scoville & Milner, 1957; Wilson & Baddeley, 1988*). *Ranganath and Blumenfeld (2005)* expressed concerns over whether the patient data really do provide clear evidence for distinct stores. They drew attention to the fact that LTM patients with medial temporal lobe damage have difficulties with some STM tasks, and suggest that this "questions whether theories of memory need to propose neurally distinct stores for short- and long-term retention" (p. 374). However, in a later review of the literature, *Jeneson and Squire (2012)* pointed out that the impairment of LTM patients in STM tasks is only observed with supraspan stimuli that exceed the capacity of STM. Under these circumstances even neurologically normal individuals will necessarily have to rely on a combination of STM and LTM, so it should be no surprise that LTM patients have problems. Furthermore, the finding that LTM patients are sometimes impaired in STM tasks is fully consistent with the two-store view; short-term recall can clearly benefit from a contribution from LTM, even if the two are functionally separate. For example, according to redintegration accounts (*Gathercole, Pickering, Hall, & Peaker, 2001; Hulme et al., 1997; Lewandowsky & Farrell, 2000; Saint-Aubin & Poirier, 1999a; Schweickert, 1993; Schweickert, Chen, & Poirier, 1999*), information in LTM can be used to reconstruct degraded traces in an STS at recall. Thus, even though LTM may play no role in the maintenance of information in STS, it can potentially improve performance in STM tasks by aiding retrieval.

The most powerful behavioral evidence for a form of distinct short-term storage comes from the early work of *Baddeley and colleagues (Baddeley, 1966a, 1966b; Baddeley & Ecob, 1970)*, which led to the development of the concept of a phonological store. The phonological store is assumed to hold speech-based representations and to have duration of just a few seconds. Information within the store can be refreshed by an articulatory control process (rehearsal) so as to prevent decay, and that same process can be used to recode visual information into a phonological form (*Baddeley, Lewis, & Vallar, 1984*). Together these form the phonological loop component of the Working Memory model (*Figure 1B*).

The critical evidence for a phonological store with a limited duration is that while memory confusions at short retention intervals are primarily phonological in nature, confusions at longer retention intervals tend to be semantic. In immediate serial recall, phonological similarity impairs order recall and can even improve item recall (*Fallon, Groves, & Tehan, 1999*) but semantic similarity has no impact on order recall (*Saint-Aubin & Poirier, 1999b; Saint-Aubin & Poirier, 2000*). Furthermore, when the items to be remembered are presented visually as opposed to auditorily, the phonological confusions can be eliminated if participants are re-

quired to perform articulatory suppression during presentation (Baddeley et al., 1984; Murray, 1968; Peterson & Johnson, 1971). Articulatory suppression is assumed to occupy the articulatory control process and prevent it from being able to recode the visual input into a phonological form.

These findings provided the motivation for separate phonological and visual short-term stores in Baddeley and Hitch's (1974) Working Memory (WM) model. More recently Baddeley (2000) has expanded the working memory framework to include a component called the episodic buffer (Figure 1B). The episodic buffer is responsible for temporary storage beyond the capacity of the model's short-term storage systems, and for binding representations across different modalities. One line of evidence supporting the need to add yet another memory system comes from studies of prose recall. For example, the STM patient PV had an auditory word span of one but could remember up to five words when they formed a meaningful sentence (Vallar & Baddeley, 1984).

The strongest evidence for a separate STS does not require carefully controlled laboratory experiments at all. The mere fact that we can repeat nonsense words like 'blontestaping' that we have never heard before shows that memory cannot be based simply on activating preexisting representations in LTM. Although there may well be preexisting representations of phonemes, features or syllables, these are not sufficient to construct a mnemonic trace of a previously unencountered nonsense word.

Short-Term Memory and Working Memory

The terms STM and working memory (WM) are often used interchangeably, and also inconsistently. In this article the primary focus is on whether there are distinct short- and long-term stores. Working memory is usually thought of as a much broader concept, often encompassing processing as well as storage. For example, both the Baddeley and Hitch (1974) and the Cowan (1988) WM models include a central executive component which can operate on the contents of STM (see figure). In the Baddeley and Hitch model, modality-specific short-term stores form part of the overall WM system. The distinguishing feature of WM is that it provides a 'mental workspace' (Logie, 2003) that can hold information in a temporary form that can be manipulated and updated. Although there are different views on exactly what constitutes WM (for review see: Aben, Stapert, & Blokland, 2012), the common feature of these views is that rather than being simply a passive store, WM is a system that allows information to be actively manipulated. For example, in the context of visual Luck and Vogel (2013) defined WM "as the active maintenance of visual information to serve the needs of ongoing tasks" (p. 392). Baddeley (1992) described WM as a "system that provides temporary storage and manipulation of the information necessary for such complex cognitive tasks as language comprehension, learning, and reasoning." (p. 556).

Although this distinction between passive stores and an active mental workspace is widely observed, some authors also use the broader term WM when describing studies that involve only passive storage. This is particularly so in the case of work on visual STM. In contrast, Nee and Jonides (2013) use the term STM to refer to all memory over the short term, whether active or passive. This includes the 'activated LTM' component of Oberauer's (2002, 2009) three-state model. One would more commonly expect this broader conception to be referred to as WM. Ericsson and

Kintsch (1995) add further complexity by proposing that there is a form of long-term WM. Although these terminological inconsistencies can lead to confusion, our concern here is with the nature of the storage mechanisms themselves rather than the higher level cognitive processes that might operate of the contents of such stores. Nevertheless, it is worth pointing out that WM tasks like reasoning or language comprehension will necessarily recruit wide-ranging neural and cognitive processes that will simultaneously engage both short- and long-term storage systems.

Interactions Between STM and LTM

Although multistore models have distinct short- and long-term stores, this does not imply that there are distinct short- and long-term tasks that tap exclusively into short- or only long-term memory. Atkinson and Shiffrin (1968) wrote that "According to our general theory, both STS and LTS are active in both STS and LTS experiments" (p. 101). There is no particular point in time that would mark a transition between tasks that only involve an STS and those that involve only an LTS. Even under the view that there are separate short and long-term storage systems, the two systems should operate in concert. Interactions between LTM and STM have already been mentioned in the context of patients with deficits in STM or LTM. Long-term learning of novel words or digit sequences depends on STM, and performance in STM tasks is influenced by information in LTM. For example, although STM is predominantly phonological, performance in STM tasks is nevertheless influenced by lexical or semantic factors, or by other information stored in LTM. For example, words are recalled better than nonwords (Brener, 1940; Hulme, Maughan, & Brown, 1991), high-frequency words are recalled better than low-frequency words (Gregg, Freedman, & Smith, 1989; Hulme et al., 1997; Watkins & Watkins, 1977), concrete words are recalled better than abstract words (Bourassa & Besner, 1994; Walker & Hulme, 1999), sentences are recalled better than random word lists (Brener, 1940), sequences of letters are recalled better when they have a closer approximation to the statistics of English (Baddeley, 1964), and lists constructed from familiar sequences are recalled better than lists composed of unfamiliar sequences (Botvinick, 2005; Botvinick & Bylsma, 2005; Mathy & Feldman, 2012).

Some have interpreted this as evidence for a version of the activation view in which STM depends on activation in LTM language networks (Acheson, MacDonald, & Postle, 2011; Poirier, Saint-Aubin, Mair, Tehan, & Tolan, 2015). An alternative view is that although there is a separate STS, long-term semantic memory is activated in STM tasks, and the sustained semantic activation is fed back to STM to aid retention (Acheson, MacDonald, et al., 2011; Huttenlocher & Newcombe, 1976; Jefferies, Frankish, & Lambon Ralph, 2006a, 2006b; Jefferies, Frankish, & Noble, 2009; Patterson, Graham, & Hodges, 1994; Poirier et al., 2015; Savill, Metcalfe, Ellis, & Jefferies, 2015; Stuart & Hulme, 2000). This is sometimes referred to as semantic binding. Both of these possibilities run into exactly the same problems as the view that STM is activated LTM. Although almost all of the data comes from serial recall tasks, there is no indication of how simply activating semantic memory could code serial order.

The alternative view is that any benefit conferred by LTM occurs only at recall and operates by means of the kind of redintegration process described earlier (Schweickert, 1993). Redinte-

gration is most commonly thought of as operating at the level of individual items. For example, a degraded mnemonic trace might be more likely to be reconstructed as a high-frequency word than as a low frequency word. Consistent with this view, factors such as lexicality, word frequency, and semantic information affect recall of item information but not order information (e.g., Poirier & Saint-Aubin, 1996; Saint-Aubin & Poirier, 1999b). However, Acheson, MacDonald, et al. (2011) reported a semantic effect on order recall. They combined a serial recall task with a secondary task requiring participants to make either semantic or orientation judgments about pictures. The semantic tasks led to more order errors than the orientation task, and produced more errors for word list than for nonword lists. They concluded that the semantic task interfered with the temporary activation of semantic representations within the language production architecture that could support STM. According to this view, nonwords would be unaffected by the nature of the secondary task as they have no semantic representations. Acheson et al. acknowledge that it may well be possible to explain their results in terms of reintegration, but they suggest that reintegration should operate at the item level and not benefit memory for order. However, the Bayesian framework presented by Botvinick (2005) was specifically designed to explain how information in LTM could help recover order information. In this framework, any information available from LTM, whether general LTM or a specifically linguistic memory, could potentially influence retrieval of both item and order information. Therefore it is still possible that their results could be explained by reintegration. From a psycholinguistic perspective, it would seem that listeners must necessarily form an integrated memory representation that combines phonological, syntactic and semantic information, although establishing exactly how those representations are stored and interact with each other remains a challenge.

A further form of interaction between LTM and STM happens at encoding. In the Baddeley and Hitch working memory framework, visually presented words must be recoded into a phonological form in order to be stored in the phonological store, and this recoding process must involve contact with representations in LTM to generate the necessary phonology. As least in the case of verbal STM, maintenance of information is helped by rehearsal which involves recycling information in STS. At each phase in the recycling process there is an opportunity for LTM to be involved in the reconstruction of degraded traces in STS. Furthermore, if items that are being rehearsed are thought of as being in the focus of attention, they must inevitably make contact with the same LTM representations that would be involved in the perception of those items.

Finally, long-term learning is not restricted to situations where there is an explicit intention to learn. Implicit learning occurs even in tasks which ostensibly only require STM. Long-term learning, and hence, access to LTM, must take place continuously if we are to be able to learn about the relationships between events occurring too far apart to be held within STM. For example, in the classic Hebb (Hebb, 1961) learning paradigm, performance in serial recall of lists of verbal material improves over successive repetitions. The implication of this is that LTM must be continuously engaged even in what might appear to be 'pure' STM tasks, so that some long-term learning can take place on the very first encounter with a list. If it did not, then every occurrence of the list would

effectively be treated in exactly the same way as the first encounter and learning could never begin.

Taking these factors together, it is very difficult to imagine any set of circumstances in which LTM will not be involved to some degree in what might seem to be a pure STM task. As we will see below, the fact that LTM cannot simply be 'turned off' in STM experiments has implications for the interpretation of neuroimaging studies that have been presented as evidence against multistore models.

As STM and LTM have to perform similar functions, they are likely to share similar features. In particular, both involve basic processes such as encoding, retention, and recall. Both are subject to loss of information through decay or interference, and both will be impaired when items to be remembered are similar. Brown, Neath, and Chater's (2007) SIMPLE (scale invariant memory, perception and learning) model emphasizes these shared properties. Their model illustrates how a general principle of scale-independent similarity can be applied to simulate a range of key empirical findings in both STM and LTM without the need to make any distinction between STS and LTS. However, by focusing on the commonalities their model is unable to account for critical data such as the neuropsychological dissociations seen in STM patients or the change in the representational basis between STM and LTM. The success of the model is attributable to the fact that it embodies broad principles that apply to memory and perception quite generally. If these principles are indeed general they would be expected to apply equally well to separate short- and long-term stores. There is no contradiction between general principles and separate stores.

Cognitive models of STM as activated LTM. Atkinson and Shiffrin's, 1971 paper is considered to be the classic statement of a two-store model. Figure 1A, based on the first figure in their paper, represents STS and LTS as separate boxes. Despite this they also say that "One might consider the short-term store simply as being temporary activation of some portion of the long-term store" (p. 83). More directly, Shiffrin (1975) wrote that "STS is the activated subset of LTS" (p. 214). Similarly, Norman (1968) talks about primary and secondary storage but also states that they "are different aspects of one large storage system" (p. 535). However, the theoretical force of these statements is unclear. Nothing seems to rest on this assumption and there is no discussion of its motivation or theoretical implications. This idea was developed further by Cowan (1988) who is regularly cited as supporting the idea that STM is activated LTM. Cowan stated quite clearly that "short-term storage should be viewed as an activated subset of long-term memory" (Cowan, 1988, p. 185), where short-term storage includes both the phonological store, and the visuospatial sketch pad. Additionally, in his model, a subset of activated LTM corresponds to the focus of attention (Figure 1C). However, on closer reading, much of what Cowan says has close parallels with the Baddeley and Hitch multistore model.

In later writings, Cowan and Chen (2008) state that "We thus propose that although the mechanisms of short-term memory are separate from those of long-term memory, they are closely related" (p. 104). They also acknowledge that "A phonologically based storage and rehearsal mechanism, such as the phonological loop mechanism of Baddeley (1986), may come into play primarily when items have to be recalled in the correct serial order" (p. 94) and "The phonological storage mechanism may be another in-

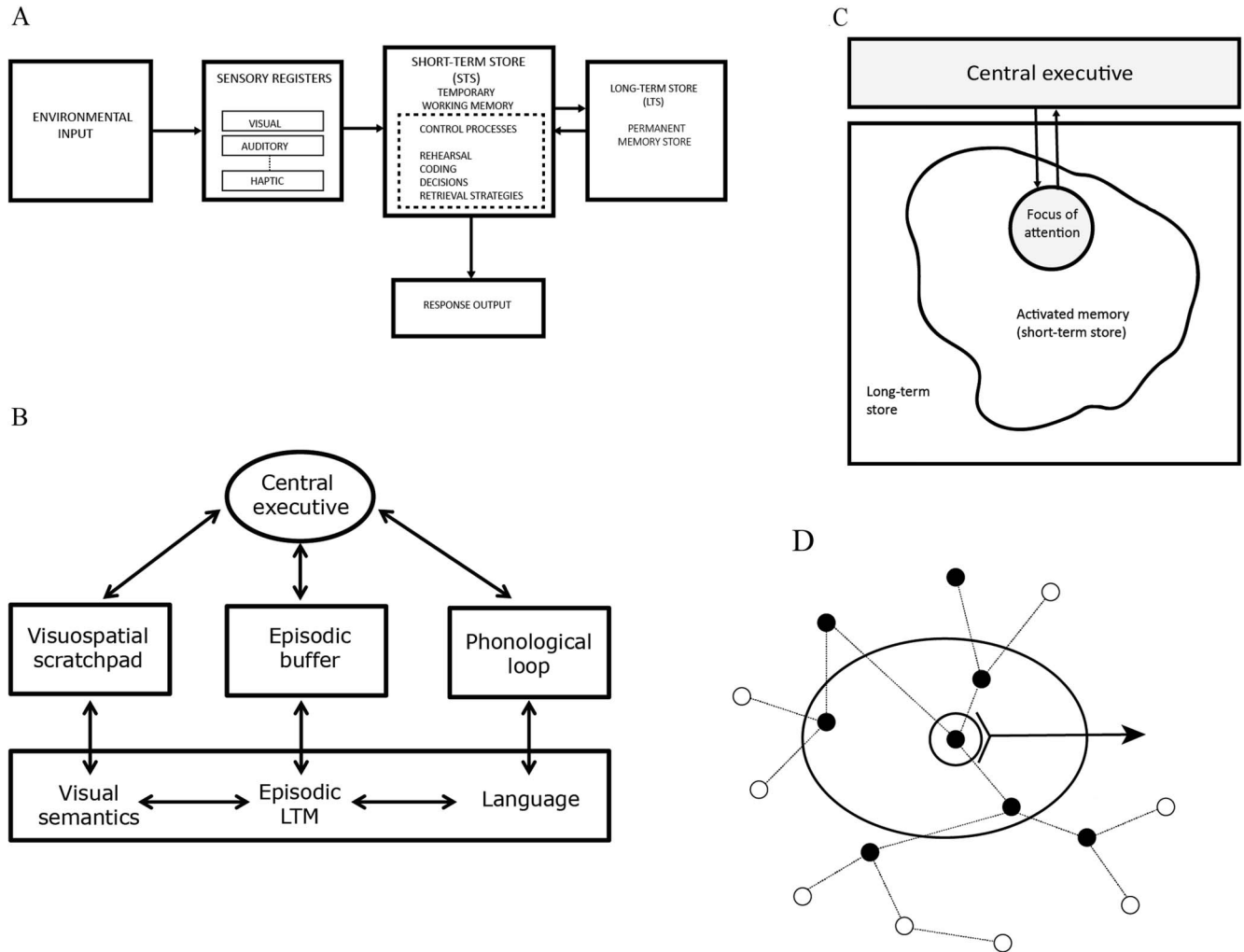


Figure 1. (A) Atkinson and Shiffrin's (1971) model of STM. (B) Baddeley and Hitch's working memory model (Baddeley, 2000). (C) Cowan's (1988) model of memory and attention. (D) Oberauer's (2002) model of working memory. Nodes and lines correspond to long-term memory representations. Black nodes are activated. Nodes within the large oval are in the direct access region. One node is in the focus of attention.

stance of the use of the focus of attention to assemble a new structure in long-term memory" (p. 95). Cowan and Chen also appreciate that some extra mechanism is required to store novel information, such as serial order, that may not already be represented in LTM. In common with the WM model, Cowan's model has a central executive which, among other things, is involved in "the maintenance of information in short-term memory through various types of rehearsal" (Cowan, 1988, p. 171). Cowan (2008) wrote that "Baddeley's (2000) episodic buffer is possibly the same as the information saved in Cowan's focus of attention, or at least is a closely similar concept." At the very least then, Cowan's model has the following components: LTM, an episodic buffer, a phonological store, and a rehearsal process controlled by a central executive. These would seem to correspond very closely to the components of Baddeley and Hitch's (1974) multistore WM model. Given that not all short-term storage in Cowan's model is supported solely by activated LTM, the crucial question then is

what is the remaining force of the claim that STM is activated LTM? Is there any part of the process of retaining information over the short-term that can be served simply by activating LTM? One factor that makes it hard to answer this question is the absence of a computational specification of what it means for LTM to be activated, and of how that activation then supports memory. As Cowan's position has evolved to accommodate a broader range of behavioral data, it has had to respect the fact that very little of that data can be explained purely in terms of activation. His theoretical position no longer equates STM directly with activated LTM.

Oberauer (2002, 2009) has proposed a three-state model which also gives a central role to activated LTM (Figure 1D). He makes an additional distinction between activated LTM and a direct access region. A similar three-state model has been proposed by Nee and Jonides (2013). Oberauer views LTM as an associative network in which related items can automatically activate each other. A subset of those activated items corresponds to a region of

direct access, and items in the direct access region can be accessed immediately. Within the direct access region a smaller set of items is in the focus of attention. It is not clear whether the direct access region should be assumed to be a region of LTM. I will return to this question later.

Is there neuroimaging evidence for STM as activated LTM?

Historically, much of the debate as to whether there are separate long- and short-term stores has focused on whether various manipulations have different effects on long-term and short-term retention. For example, are there differences in the form of the serial position curve observed over different retention intervals? Wickelgren (1973) reviewed the behavioral data available in the early 1970s and argued that although the evidence favored the two-store account “the vast majority of all the cited evidence is worthless for making the distinction between short and long traces” (p. 426). Very little has changed since then. But might neuroimaging data be able to provide insights beyond those that can be had from behavioral data? If it could be shown that brain regions known to be responsible for LTM are activated during STM tasks, surely this would be evidence that those LTM regions were supporting STM? This unsound line of reasoning provides the rationale for the design and interpretation of a number of neuroimaging studies which have been presented as providing evidence that STM is simply activated LTM.

A flavor of the claims made in some neuroimaging papers can be found in the following quotations: “This result implies that activated long-term memory provides a representational basis for semantic verbal short-term memory, and hence supports theories that postulate that short- and long-term stores are not separate” (Cameron, Haarmann, Grafman, & Ruchkin, 2005, p. 643, abstract). “Temporary activation of long-term memory supports working memory” (Lewis-Peacock and Postle, 2008, title). “these findings support single-store accounts that assume there are similar operating principles across WM and LTM representations, and the focus of attention is limited.” (Öztekin, Davachi, & McElree, 2010, abstract). “Thus, the theoretical claim of our model—that maintenance in VWM entails the activation of long-term, speech-based representations—invalidates a central tenet of the multi-component model,” (Acheson, Hamidi, Binder, & Postle, 2011, p. 1359). Note that the first three quotations come from either the abstracts or the title of the papers so are presumably central to the claims being made by the authors.

Exactly what is the evidence that supports these forceful claims? Öztekin et al. measured hippocampal activity during a 12-item probe-recognition task and found that “activation of the medial temporal lobe was observed for all serial positions other than the last position, and that the activation level could be predicted from the underlying memory strength.” (Abstract). More generally, hippocampal activation decreased over item positions and this correlated with recognition accuracy. On this basis they claim that their “results are inconsistent with the traditional view that posits that there are distinct stores for WM and LTM” (p. 1130). This conclusion is based on what (Baddeley, 2003) has called the correlational fallacy; the assumption “that any variable that is broadly correlated with performance on the paradigm is crucial to it” (p. 729). As has already been shown, hippocampal activity is to be expected in STM tasks simply because LTM cannot be turned off in cases where it might not be needed. In fact LTM is very likely to be needed in this study, as a 12-item list will be well beyond the

normal span of STM. Öztekin et al. favor an explanation based on Oberauer’s (2002, 2009) three-state memory model which posits that memory items can be in three states: a single item in the focus of attention, three or four items that are in an activated state in LTM but need to be retrieved from working memory, and items in a passive state which need to be retrieved from LTM. The first and last of these states are common to multicomponent models; one or two items may be in the focus of attention and some, especially in a 12-item list, will need to be retrieved from LTM (or possibly the episodic buffer component proposed by Baddeley, 2000). The critical question is whether these data provide any evidence that some items are in an activated state in LTM. As hippocampal activation is expected as a matter of course, even under the assumption that there is a separate STS, these data do not provide an answer.

Cameron et al. (2005) based their claim that activated LTM provides the representational basis for STM on data from an ERP study. They compared two conditions, one of which was assumed to involve memory for semantic information and the other was intended as a control task. In the memory task, participants saw a sequence of three words presented at a rate of one every 1000 ms, followed after 4000 ms delay period by a probe word. The participants’ task was to indicate whether the probe word was related to one of the three preceding words. In the control task, participants were required to indicate whether one of the words was unrelated to the other two, so in this case there was no memory load during the delay period. In both cases an incidental probe word was presented auditorily 2000 ms after the beginning of the delay period and this probe could be either semantically related or unrelated to the third word. The ERP analysis found an effect of semantic relatedness that was greater for the memory task than the control task, and this led the authors to conclude that STM was supported by LTM. This is a further example of the correlational fallacy. What the data show is a priming effect which is larger in the semantic memory task than in the control task. There is no evidence that whatever causes the semantic priming also plays a causal role in supporting STM. In the memory task participants are likely to be recirculating the to-be-remembered items. As items come to the fore they will be expected to elicit similar perceptual and semantic processes to those elicited at encoding.

Lewis-Peacock and Postle (2008) used fMRI combined with multivariate pattern analysis to establish whether a classifier trained to discriminate between three categories of stimulus (people, objects, locations) in a judgment task could classify neural activity patterns during the delay phase of an STM task. The judgment task involved answering questions such as how often the participants encounter a particular object. The classifier was trained on the data from this judgment task. In a second phase of the experiment participants learned to associate pairs of the items. In the STM task itself, participants were presented with a target stimulus for 1 s and, 11 s later, they saw a probe stimulus. Their task was to judge whether the probe stimulus was the associate of the target. During the delay period the classifier trained on the LTM judgment task could decode both the category of the target and the category of the probe above chance. Lewis-Peacock and Postle argued that “successful decoding of working memory activity would provide conclusive evidence for the activated LTM model” (p. 8768).

This is a further example of the correlational fallacy. The fact that LTM activity can be decoded during short-term retention interval does not imply that those LTM representations are responsible for short-term retention. Furthermore, the classifier was trained only to distinguish between representations of different semantic categories. It is quite possible that memory was supported primarily by some form of phonological storage in an STS, but Lewis-Peacock et al. only looked for semantic activation. Perhaps acknowledging this, Lewis-Peacock and Postle concede that “This model does not necessarily rule out a parallel contribution to delay-task performance by specialized working memory systems” (p. 8673). Data from a study by [Shivde and Thompson-Schill \(2004\)](#) suggests that this might well be the case. They used a similar STM paradigm where the task was either to judge whether the probe word was a synonym of the target or shared a vowel sound with a nonsense word probe. They found that the semantic task elicited delay period activation in bilateral inferior frontal gyrus and left middle temporal gyrus, while the phonological task elicited activation in left superior parietal cortex. In other words, when the task was phonological, the activation was in a region more commonly associated with phonological STM. Shivde and Thompson-Schill used univariate analyses so it is impossible to say whether a classifier trained on a phonological contrast would have been successful in decoding phonological information in the retention interval of their phonological task, or even possibly their semantic task. Not all brain regions that are activated will contain representations detectable by a classifier, and classification can often be successful in regions that show no univariate activation (e.g., [Harrison & Tong, 2009](#)). However, these data clearly show that the region that is activated during retention depends on what is being retained. There is more to the story than the simple claim that STM is activated LTM. Finally, a task requiring retention of a single item is one of the few that that could conceivably be performed solely by activating that item in LTM. The task demands might be too simple to be able to guarantee that a dedicated STS would be fully engaged.

[Acheson et al. \(2011\)](#) claim that “WM reflects the temporary activation of LTM representations” (p. 1359). They argue that “the theoretical claim of our model—that maintenance in VWM entails the activation of long-term, speech-based representations—invalidates a central tenet of the multi-component model” (p. 1359). However, their study really focused on a different and less contentious issue, which is whether short-term storage of verbal material depends on the same processes that are responsible for perception or production. Even adherents of the multicomponent view have considered the possibility that the phonological store might be isomorphic with the memory system supporting speech production ([Page, Madge, Cumming, & Norris, 2007](#); [Page & Norris, 1998a](#)). [Page and Norris \(1998a\)](#) presented a connectionist implementation of their Primacy model of immediate serial recall which had close parallels to the speech production models of [Dell and O’Seaghdha \(1992\)](#) and [Levelt \(1992\)](#). [Acheson et al.](#) examined immediate serial recall of five-item lists using rTMS over either the MTG or posterior STG during a 3-s retention interval. MTG and posterior STG were chosen as they are implicated as the sites of two stages of production planning ([Indefrey & Levelt, 2004](#)). Their rationale was that if rTMS over these production areas impaired recall then these areas must be involved in STM. Despite the strong claims made on the basis of this data, the results

are rather weak. rTMS had no significant effect on item order errors, and an influence on omissions only in posterior STG. Let’s pretend that there was a large effect of rTMS on both item and order errors in both regions. What might we conclude? At the core of the Working Memory model is a speech-based articulatory rehearsal process. Unless participants are actively prevented from doing so, they will almost certainly be trying to rehearse the five-item list during the 3s retention interval. rTMS that interfered with production should also interfere with rehearsal. The only surprising finding here is that the effects were not larger than they actually were.

A positive feature of the [Acheson et al.](#) study is that they used a standard serial recall task. Serial recall is the most widely used task to study short-term verbal memory and is the primary target for computational models of verbal STM. The other neuroimaging studies described here used nonstandard tasks that do not have a well-established relation to specific components of memory in multistore models. For example, [Lewis-Peacock and Postle](#) used a retention interval of 11 s, which is well beyond the duration of the phonological store. Indeed, this seems to be a common problem. [Talmi, Grady, Goshen-Gottstein, and Moscovitch \(2005\)](#) tried to make a case against the multistore models using data from a study where participants had to remember lists of 12 words. This is twice the normal span for words. The choice of task may in part be determined by the difficulties associated with using standard laboratory tasks in a neuroimaging environment. However, it is technically possible to run standard immediate-serial recall experiments using fMRI (e.g., [Kalm, Davis, & Norris, 2013](#)).

[Öztekin et al.](#), [Cameron et al.](#), and [Lewis-Peacock et al.](#) have all made the mistake of assuming that the presence of neural activation of brain regions normally associated with LTM during an STM task implies that long-term representations in those regions form the causal basis of STM. They have all committed the correlational fallacy and failed to appreciate [Atkinson and Shiffrin’s](#) point that “both STS and LTS are active in both STS and LTS experiments.” The only way to construct a model of STM where there would be no LTM activation in a STM experiment would be by assuming that information in STS was stored in opaque containers such that the contents of that store were inaccessible to other processes. If information in STS is visible to other processes then it should activate those processes too.

What is activation? If STM is supported exclusively by activated LTM, it seems reasonable to ask what computational function is performed by activation that would enable it to encode, maintain, and retrieve information from STM. This is a fundamental and largely unrecognized problem with all models invoking activated LTM. Although the core explanatory concept is activation, there is no explicit definition of what activation means. In the memory literature the term activation often refers to the deployment of a limited capacity resource that can be used to support WM ([Anderson, 1983](#); [Cantor & Engle, 1993](#); [Just & Carpenter, 1992](#)). However, there is no computational definition of activation of LTM that would explain how that ‘activation’ might be sufficient to maintain representations in STM. As [Phillips \(2003\)](#) noted in his commentary on [Ruchkin et al.’s \(2003\)](#) position paper, sometimes activation was used “in a psychological sense, sometimes to refer to EEG measures, and sometimes to refer to underlying neuronal activity” (p. 752). Worse still, in the absence of an explicit computational model there is no agreed psychological

sense of *activation* to fall back on; it simply is not at all clear what it means to say that STM might be activated LTM. Without an understanding of the computational functions served by the neural activation seen in LTM systems, any attempt to interpret correlated activity in brain regions associated with LTM during short-term retention remains highly speculative. We know too little about how to interpret neural activation for it to form the basis of an explanation. Even if we knew which neural processes caused activation, this would not tell us what computational function was being performed by those processes.

A further problem for activation-based theories is that recent neuroimaging work suggests that storing information in STM may not depend on sustained neural activation of any form. Several studies (LaRocque & Lewis-Peacock, 2013; LaRocque, Riggall, Emrich, & Postle, 2016; Lewis-Peacock, Drysdale, Oberauer, & Postle, 2012; Sprague, Ester, & Serences, 2016) have tried to address the question of whether STM is supported by sustained neural activation. These studies all used a retrocuing paradigm combined with multivariate pattern classification. In an initial phase of the experiments a classifier was trained to discriminate between different stimuli that would have to be remembered. For example, in the LaRocque et al. (2016) study participants had to remember the direction of motion of arrays of dots. The pattern classifier was then used to track the progress of activated memory representations over the course of a retention interval while participants shifted their attention between one of two stimuli that might have to be remembered. In the critical conditions, two visual items are presented along with a cue indicating which item to attend to. There is then a delay period of 16s. Halfway through the delay period a retrocue is presented to indicate whether the participant should maintain attention on the same attended item or switch attention to the previously unattended item. A test probe is then presented at the end of the delay period. Performance on the test probe is high regardless of which item is probed. All of the studies reported that during the first half of the retention interval, only the item that was attended to could reliably be decoded by the pattern classifier. However, when the retrocue indicated that attention should be switched to the previously unattended item, then that item could now be decoded. That is, only attended items could be decoded but it did not matter whether those items had previously been in an unattended state where their representations could not be decoded. LaRocque, and colleagues (LaRocque & Lewis-Peacock, 2013; LaRocque et al., 2016; Lewis-Peacock et al., 2012) concluded that only items that were in the focus of attention were activated, and that sustained activation was not necessary to maintain items in STM. Sprague et al. describe the unattended information as being stored in latent or activity-silent codes. In other words, not only is there no evidence that STM involves sustained LTM activation, but it may not require sustained activation of any form at all over intervals of up to 16s. In reviewing some of this data LaRocque, Lewis-Peacock, and Postle (2014), two of whom had previously argued that STM was activated LTM (Lewis-Peacock & Postle, 2008), quite reasonably pointed out that “that this result should be taken as a call for clarification of what is meant by “activated.” . . . Rather than referring to such items as being retained in the “activated” portion of LTM, a word more in keeping with the function of STM—such as “accessible”—might be more appropriate to describe STM items outside the FoA” (Focus of Attention. p. 9). However, the shift from ‘activated’ to

‘accessible’ completely eliminates the theoretical force of any claim about the role of LTM in supporting STM. Without further clarification, the term ‘accessible’ does little more than redescribe the data.

Part of the problem with explanations based on activation may stem from the tempting allure of the activation metaphor. Activation has been a popular theoretical metaphor in psychology since the 1960s. In the context of memory models, activation usually refers to the strength of a representation, and activation levels are inferred from how well an item is remembered. With the rise of connectionist or neural network models in the 1980s, activation became even more central to psychological thinking. Activation is the currency of communication between the artificial neurons. Connectionist models had the great advantage of producing quantitative simulations that could be directly tested against the data. The concept of activation in these models has clear computational underpinnings. However, it sometimes seems that the successful use of activation in connectionist modeling has been taken to license indiscriminate use of the term in verbal models.

Computational models. In any scientific endeavor, progress depends on developing theories that make clear and testable predictions. In STM research much of that progress has been driven by computational modeling. For example, the basic principles of Baddeley and Hitch’s WM model have been formalized in computational models of the phonological store (Burgess & Hitch, 1992, 1999, 2006; Page & Norris, 1998b). These and other computational models of STM (e.g., Botvinick & Plaut, 2006; Brown et al., 2000; Farrell, 2012; Farrell & Lewandowsky, 2002; Henson, 1998) produce accurate quantitative simulations of most of the benchmark data from STM tasks: the form of the serial position curve, the nature of the errors in serial recall, effects of perceptual grouping, complex patterns of phonological similarity effects, sequence learning, and much more. One could argue that the problem with these models is that they are so good that it is becoming increasingly difficult to devise experiments to distinguish between them. Nevertheless, these are all concrete models that can be subject to empirical scrutiny. In contrast, there are no computational models based on activation of LTM. Before activation-based models can be taken seriously they need to be clearly formulated and to be shown to be competitive with existing computational models.

STM must be able to represent multiple tokens. The spoken sentence “Buffalo buffalo buffalo Buffalo buffalo” contains five tokens of the same phonological word form – ‘bʌfələʊ. If you do not know that a buffalo is a North American bison, that ‘to buffalo’ means ‘to confuse,’ and that there is a town called Buffalo, this sentence will probably leave you completely buffaloed.¹ In fact, even if you know all three of those it’s still hard to understand. Nevertheless, if you hear this sentence it is very easy to repeat it back—it’s just the same word repeated five times. If STM were nothing more than the activated portion of LTM then a sentence or list containing five repetitions of the same phonological word form

¹ The sentence should be interpreted as indicating that bison who are from the town of Buffalo confuse (buffalo) other bison from the town of Buffalo. It is commonly used as an example of several different linguistic phenomena and may have been generated independently by a number of researchers. See <http://www.cse.buffalo.edu/~rapaport/buffalobuffalo.html>

could not be recalled. There would just be a single very strongly activated representation of $\beta\lambda f\alpha l\epsilon\upsilon$.

This is an example of what Jackendoff (2002) called ‘the problem of two’ – the cognitive system needs to be able to represent multiple tokens of the same type.² Of course, it might be the case that LTM contains multiple representations of the word ‘buffalo’ that can each be activated separately. But how many? Is it possible to repeat the sentence “Buffalo buffalo Buffalo buffalo buffalo buffalo Buffalo buffalo”?³ Maybe LTM contains a whole herd of buffalo? Possibly these sentences could be coded by a single LTM representation with either five or eight units of activation. But what about a list like “buffalo cowboy cowboy buffalo”? Could activation levels reliably differentiate between the lists “buffalo cowboy cowboy buffalo” and “cowboy buffalo buffalo cowboy” so that they could be recalled in the correct order? The idea that STM is nothing more than activated LTM fails to provide a solution to ‘the problem of two.’ Even the simple requirement to remember sentences or lists containing two occurrences of the same type (word) cannot be supported solely by activating LTM. In the memory literature the importance of being able to represent multiple tokens was highlighted by Potter (1993) in discussing what she called “Very short-term conceptual memory.” She pointed out that “It is not enough just to activate the type representation of each word in LTM; one must set up an episodic representation that includes a token of each word type, tokens that are pointers to or copies of the long-term representation of the words and their meanings” (p. 159). Note that the need to represent multiple tokens is not limited to verbal processing. In the context of visual perception and attention Kahneman and Treisman (1984) made a distinction between “episodic tokens and semantic types” (p. 55) in the development of their concept of an object file. Theoretical accounts of phenomena such as repetition blindness (Kanwisher, 1987, 1991) and the attentional blink (Bowman & Wyble, 2007) also require a distinction to be made between type and token representations.

Most computational models of STM have the ability to represent multiple tokens. One straightforward way of achieving this can be found in connectionist models of immediate serial recall that are based on forming associations between item representations and representations of list position (e.g., Burgess & Hitch, 1992; Brown, Preece, & Hulme, 2000). The Burgess and Hitch model was explicitly developed as a model of the articulatory loop component of Baddeley and Hitch’s (1974) working memory model. In their model a vector representing list position slowly evolves so as to change slightly as each successive word is presented. At each position, the network learns an association between the position vector and the word. At recall, the states of the position vector can be replayed and successive states of the vector will tend to reactivate the word in the list associated with that position. Because the same word can be associated with more than one position, the model has no difficulty in remembering sequences where words are repeated. One might suggest that the words that are activated must be in LTM and this could then be construed as performing short-term storage by activating LTM. However, the crucial point is that the model’s performance depends entirely on the position-item association mechanism which exists purely to perform short-term serially ordered recall. STM is supported by a dedicated mechanism which exists in addition to LTM. That mechanism is a STM store. Furthermore, in this particular model, the memory is encoded in the weights between the

position vector and the word representations and not by sustained activation over LTM. This reminds us once again that although there is a close relationship between LTM and STM, LTM cannot do the job of supporting STM on its own. The problem of two poses an insurmountable barrier to the idea that STM might be supported simply by activating LTM. This holds regardless of the content of STS. Exactly the same argument will apply to STM for semantic information (Martin & Romani, 1994; Martin, Shelton, & Yaffee, 1994; Shivde & Anderson, 2011) or visual information.

STM must contain structured representations. The ability to represent multiple tokens is a basic prerequisite for a simple passive STM system. However, this is not sufficient for a memory store that could serve as a mental workspace for WM. A consequence of being able to actively manipulate the contents of WM is that this can create completely new structures which have never been encountered before, which can themselves be manipulated. By definition, this productivity cannot be captured simply by activating preexisting representations (Logie, 2003; Logie & Della Sala, 2003). In simple short-term verbal memory tasks the only novelty is generally in the ordering or composition of the items to be remembered. In contrast, WM has to have the capacity to store a vast and unpredictable set of representations. There might be some STM tasks that can be performed simply by activating preexisting structures, but rarely will that be possible in a WM task.

Variable binding. To hold structured representations of objects or scenes the system must also be able to perform variable binding (see van der Velde & de Kamps, 2006, for review). For example, consider how we might maintain a coherent representation of a sentence such as “The young boy saw the boy who was singing.” Here the problem is not simply representing the order of the words, or even that there are two tokens of the word boy, but appreciating that there are two different boys, one of whom is singing and one of whom is young. It’s necessary to both represent multiple tokens and the bindings between each of those tokens and other components of the sentence. However, this cannot be achieved solely by coactivating and associating existing representations (as assumed by Cowan & Chen, 2008). This can be appreciated by considering the simple case of forming associations between AB, BC, and CA. Associating these three pairs will necessarily form associations between all items in the triple ABC. It will therefore be impossible to determine whether what has been learned was AB, BC, and CA, or ABC. What is needed are additional representation of the individual conjunctions (e.g., creating an ‘AB’ node in a connectionist network). In STM this could be achieved by variable binding. That is, STM would contain a set of variables to which arbitrary representations could be assigned. For example, $x = (A, B)$, $y = (B, C)$ would construct bound representations assigned to the variables x and y . This process may or may not activate representations in LTM, but, for the reasons already given, it should not form direct associations between A and B or between B and C. This might seem to present a severe problem for Oberauer’s three state model shown in Figure 1D.

² The same argument applies even if ‘Buffalo’ and ‘buffalo’ are assumed to activate separate semantic representations and therefore correspond to different types as there will still be multiple tokens of the those types.

³ See https://en.wikipedia.org/wiki/Buffalo_buffalo_Buffalo_buffalo_buffalo_buffalo_Buffalo_buffalo for interpretation.

Oberauer depicts his model as having three concentric regions—activated LTM, the direct access region, and the focus of attention, all of which select a different subset of LTM representations (shown by the lines and nodes). However, while the diagram and the term *region* imply that the direct access region is simply a subset of activated LTM, the operations taking place within the direct access region involve much more than activation. According to Oberauer “the region of direct access is a mechanism for establishing and holding temporary bindings between contents. . . . By supporting arbitrary bindings between virtually any content with any context, this system enables the compositionality of thought that many theorists regard as a hallmark of human cognition (Fodor & Pylyshyn, 1988): we can create an unlimited number of different ideas by freely combining content elements into new structures.” (Oberauer, 2009, p. 53). The direct access region is therefore much more than just an activated *region* of LTM. It must have all of the computational machinery needed to support all aspects of higher cognition, and to have the same functionality as the central executive and episodic buffer components of the Baddeley and Hitch WM model. As with Cowan’s model, it is not clear whether activated LTM alone plays an essential part in the model’s operation, or in its ability to retain information over the short-term. The fact that Oberauer’s model needs to be supplemented with a binding mechanism reinforces the argument that a model of STM needs more than just activated LTM to make it work. Indeed, the big challenge for any model that might be based purely on activation of LTM would be to make it work.

Visual STM. The case for STM as activated LTM has been advanced predominantly in the domain of verbal STM. Perhaps this is because some verbal STM tasks might be seen as simply requiring the participant to remember a collection of familiar words, which might be achieved simply by activating existing representations. As we have seen though, the need to remember words in sequence demands that words be bound to positions. However, it has long been appreciated that perception of visual objects depends on forming structured representations that bind together features such as shape, color and location (Kahneman & Treisman, 1984; Marr, 1982; Pylyshyn, 1989, 2001). Not surprisingly, the notion that STM might simply be activated LTM has been less influential here. Combinations of visual features can produce completely novel visual objects that cannot possibly have preexisting representation in LTM. Nevertheless, configurations of complex objects can be stored in STM. Even an image as simple as an array of randomly placed identical dots cannot be represented just by activating the appropriate number of dots. The dots must each be individuated and bound to a unique spatial location. Indeed, the change detection task used in many studies of visual STM (e.g., Luck & Vogel, 1997) requires participants to judge whether an object (such as a colored square) in a particular location has changed color. There may well be other objects of the same color in other locations. The task can only be performed by binding colors to locations. The task of remembering a spatial sequence presents even more of a challenge for activation accounts. The Corsi block task (Corsi, 1973) requires participants to touch a series of blocks in the same order as the experimenter had done. Here there is no need to remember the spatial location of the blocks, as they remain in view; instead, the participant must remember the binding between the location of the blocks and their order.

In the neuroimaging literature on visual STM many of the data are consistent with what the ‘sensory recruitment’ hypothesis (Lee & Baker, 2016; Serences, Ester, Vogel, & Awh, 2009). This suggests that maintenance of information in visual STM is achieved by recruitment of the same neural mechanisms that are engaged in encoding that information. That is, information in STM is stored in perceptual areas and not in those areas involved in LTM. The sensory recruitment hypothesis is supported by data showing that information maintained in visual STM can be decoded from brain regions known to be involved in visual perception (Harrison & Tong, 2009; Serences et al., 2009). Information maintained in visual STM can also be decoded from regions of parietal cortex (Bettencourt & Xu, 2016; Christophel, Hebart, & Haynes, 2012; Sprague, Ester, & Serences, 2014; Sprague et al., 2016) similar to those that have been found to be active in verbal STM. However, at the moment it is far from clear how the neuroimaging data can be mapped onto cognitive models of visual STM. For example, some have argued that there is more than one visual STM. Basing his argument primarily on neuropsychological data Logie (2003), has suggested that there may be one store for spatial layout, and one that supports the ability to manipulate mental images, and to retain dynamic information such as sequences.

Copies or pointers? In Atkinson and Shiffrin’s modal model the store is a box that you put things in. But what do you put in the box? This was never explicitly specified. In their 1966 paper they suggested that the item “that is placed in the buffer may be considered to be an amount of information which is sufficient to allow emission of the correct response” (p. 3). One way to achieve this is to store a copy of a representation of the input. For example, in the WM model the primary responsibility for storage of verbal information falls on the articulatory loop, which stores copies of phonological representations. Alternatively, instead of storing copies, one might store pointers to representations stored elsewhere. Either would be sufficient to allow emission of the correct response.

The possibility that an STS might be based on pointers rather than copies has a number of interesting implications. If the pointers address representations in LTS then this would greatly complicate the relationship between STS and LTS. However, pointers might also operate within the bounds of a self-contained modality-specific STS. As we will see, the advantage of this option is that it can make very efficient use of the available storage capacity. Far from blurring the lines between STS and LTS, consideration of how a pointer system might work actually reinforces the case for modality-specific short-term stores.

Perhaps the simplest possible model of an STS is Conrad’s (1965) slot model—STS would contain a number of slots, and each slot would contain a copy of an item to be remembered. But what does it mean to store a copy? Exactly what is it that’s copied into memory? In principle, when trying to remember ‘dog,’ we could store a copy of everything we know about the word ‘dog.’ That might seem excessive. More plausibly, we might store only a subset of the information held in LTM; perhaps a phonological representation of ‘dog.’

An alternative to storing copies of objects is to store pointers to representations elsewhere. As every programmer knows, it is usually much more efficient to manipulate pointers than to copy entire data structures. In computing terms a pointer is the address of a

location in memory where information is stored. The great advantage of using pointers is that the pointers generally take up much less storage than the data they point to, and they eliminate the need to copy the data. For example, if the data objects were lexical representations of words, any list of words could be represented economically by a list of pointers to those objects. The STS storage requirements would remain the same no matter how complicated the lexical representations were. Importantly, STS could represent multiple tokens by having more than one pointer index the same type representation (the same word in LTS). Thus, if the capacity of STS happened to be five pointers, it would be possible to retain the identity and order of any five objects.

It is worth bearing in mind that it is possible to write two programs to perform the same computations that differ solely in terms of whether they rely on copying data or manipulating pointers. Indeed, programming languages differ in terms of whether they perform assignment by copying data or pointers. In C, “ $a = b$ ” copies the data in b to a . In Python, “ $a = b$ ” sets a to point to the same data as b . If one were to write two superficially identical programs, one in C and one in Python, they would produce the same output, but the underlying operation of the programs would be different.

Some of the confusion seen in the neuroimaging literature might come from an assumption that a distinct STS would hold representations copied from LTM, and that therefore LTM should not be activated while that information is being retained in STM. Interestingly, in computational models of STM, what is stored is generally more akin to a pointer than a copy. In the Burgess and Hitch model the weights between position vectors and word vectors act as pointers to a single representation of a word. There is no need to have multiple copies of a word.

Pointers, address space, and efficiency of storage. Although it is tempting to think that memory capacity might be determined by the number of available pointers, not all pointers are created equal. The amount of memory required to store a pointer will vary as a function of the number of objects that the pointer can potentially address. In an early 8-bit computer an 8-bit pointer could only address one of 256 memory locations. A 16-bit pointer could address one of 65536 locations. How large a pointer might be necessary in order to address any possible representation that might be stored in LTM? Let’s pretend that a 32-bit pointer would be sufficient as that could address any one of 4.3×10^9 representations (on the assumption that each representation corresponds to a single address). But an 8-bit pointer would be more than sufficient to index the set of phonemes in a language. With a memory capacity of 32 bits it would therefore be possible to store pointers to 4 phonemes, or just a single pointer to any object in LTM—which might be a phoneme. There’s a simple trade-off. For a given memory capacity it is possible to store pointers to either a large number of items selected from a small set, or a smaller number of items selected from a larger set. This has implications for how best to make efficient use of available memory capacity. Imagine that there is some set of computations in language comprehension that only requires manipulation of phonemes. If those computations can be performed in a dedicated speech store that only needs to reference speech representations, then the total memory requirement (in bits) will be much less than if those same computations have to be performed in a general purpose system that has the potential to reference all possible representations in LTM. Note

that a specialized store could be made even more efficient if it could take into account the redundancies of the language to construct a compressed representation. One estimate is that each phoneme in English phoneme could be coded by as few as 3 bits (Stilp, 2011). That is, a dedicated store might be able to store as much as 10 times as many phonemes in a given capacity than a general-purpose store. Note that the amount of data compression possible will depend how much context can be taken into account (Shannon, 1951). Nevertheless, a dedicated store that incorporated knowledge of say, syllables,⁴ or the phoneme transition probabilities of English, could make far more effective use of storage than one that had to rely on pointers to LTM. One point in favor of a phonological store like that in the WM model is therefore that it would make very efficient use of available memory capacity.

Note, however, that if we used an 8-bit pointer to represent phonemes in a phonological store, that pointer could not make direct contact with phonological representations in LTM on its own. If LTM had, say, a 32-bit address space, an 8-bit pointer could only address a subset of that space. A simple solution would be that a modality-specific store could add a bit pattern to the modality-specific pointer that would serve to select the relevant subset of LTM representations. In our simple example, the phonological store would pad out the 8-bit pointers to 32-bits with an extra 24 bits that effectively say “I’m a phonological representation” and ensure that the 8-bit pointer can address the correct part of LTM. In the brain, if phonological representations in LTM were, in a sense, stored close together, then connections from phonological STM could target specifically those representations. The same reasoning holds even if what is stored are copies. Efficiency will be greatest if only the minimum amount of information required is copied into the store. Although the pointer argument is made from a purely computational perspective, an analogous argument can be made from the perspective of neural economy (Bullmore & Sporns, 2012). If some computations involve a restricted set of processes and representations, then the neurons responsible for those computations should be strongly interconnected and physically adjacent so as to minimize expensive long-range connections.

All of these arguments will apply equally well to any system that implements the same functionality as pointers. For example, in a connectionist model relying on weighted connections between nodes, the greater the number of representations that can be stored, the more nodes and connections will be required.

Pointers and single-store models. Proponents of single store models have also invoked the idea of pointers. In arguing that working memory is a state of activated long-term memory Ruchkin et al. (2003) say that “Prefrontal cortex provides the attentional pointer system for maintaining activation in the appropriate posterior processing systems.” (abstract). That is, there is an attentional pointer system that takes care of all of the processing operations that are needed by an STS, whereas the representations that are being retained exist only in activated LTM. Ruchkin et al. do not specify exactly what is meant by an attentional pointer system, nor by activation. However, adding a pointer system to

⁴ Simply coding the input in terms of syllables rather than phonemes (allowing for the fact that there are many more syllables than phonemes) can reduce the requirements to between 3 and 4 bits per phoneme.

activated LTM at least has the potential to overcome the fundamental barrier that undermines simple activation models - the problem of two. But exactly how does such a model differ from one that posits separate short- and long-term stores? The pointer system functions as a store and must necessarily be responsible for representing multiple tokens and serial order. The claim seems to be predominantly one about neural implementation or localization rather than function; the pointers are in prefrontal cortex and the representations are in posterior systems. As I have already argued, pointers will generally provide the most efficient way to represent multiple tokens and to manipulate information in STM. However, there are also efficiency gains to be had by restricting pointers to operate on representations held within a local STS rather than using a single pointer system that can address all of LTM. However, apart from the efficiency argument, a greater concern is that, under this view, the basic behavioral phenomena that motivated the WM model remain unexplained. A pointer-based store that could address any representation in LTM would have no reason to be modality-specific - the pointers could equally well address phonological, semantic, or even visual representations. So why does verbal STM seem to be primarily based on a phonological code, and why does phonological coding of visual input appear to be eliminated by articulatory suppression? Why would the pointers point to purely phonological representations if they could equally well point to lexical representations?

Ruchkin et al. say that “findings of multiple, modality-specialized short-term stores are fully compatible with the position that short-term memory corresponds to activated long-term memory representations, given that long-term memory involves multiple, modality-specialized stores.” (p. 712). However, the behavioral data show that the qualitative behavior of the different stores is different. At the very least they have different time-courses. If short-term retention is controlled by a single pointer system, then all short-term storage should have the same time-course.

Most of the data that Ruchkin et al. marshal in support of their view come from electrophysiology. However, in common with the neuroimaging data reviewed above, this is entirely correlational and shows only that, during STM tasks, there is neural activity in brain regions usually associated with LTM or perceptual processes. We have already seen that, according to a multistore view, LTM and perceptual regions should also be activated, so finding correlated activation does not tell us whether that activation plays a causal role in retention. The only data we have that speaks directly to that question is the data from STM patients.

Ruchkin et al.’s position makes it clear that we should not view the question of whether there are distinct long-term and short-term stores as being a simple dichotomy. There is a very large space of possible models. The simplest possible model might seem to be Atkinson and Shiffrin’s modal model, but even there the STS could contain copies, or could contain pointers. In most cases these two alternatives would produce indistinguishable behavior. Pointers might address representations in LTM, or representation in modality-specific stores. A short-term phonological store might operate with pointers to representations held within that store itself. Additionally, it might be a specifically mnemonic store, or parasitic on processes responsible for perception or speech production.

An additional layer of complexity arises when we consider how any of these memory systems might be implemented in the brain.

For example, if there is an STS containing copies of information in LTM, will the binding between the two sets of representations elicit neural activity? Or, to put it the other way around, what does neural activity tell us about the computations involved? More fundamentally, what is the mechanism of neural binding underlying STM? The space of possibilities grows even larger when we move from just considering STM to thinking about working memory and how to manipulate representations and coordinate them across different modalities. Note that I have referred to a system holding pointers to LTM as an STS even though the representations are in LTM. If there is a system where the STS contains pointers to LTM, should we really call this a STS, or is it just a pointer system? My own inclination is to stick with the term STS, as the pointers are doing all of the hard work. They also control the capacity of STM—more pointers means greater capacity.

Summary. When reading many of the papers cited here it is easy to form the impression that there is a clearly formulated model of memory in which STM is simply the activated portion of LTM. No such theory exists. This should not really come as a surprise as such a model would founder on the problem of two. Any model of STM must have some additional mechanisms beyond simple activation. There must be mechanisms to control the activation, to store multiple tokens, to recall those tokens in the correct order, and to create novel structured representations that cannot yet be in LTM. Only then could such a model begin to compete with existing computational models of STM developed within a multistore framework. However, those additional mechanisms would effectively amount to a dedicated STS. In a system relying on storing pointers in a STS, STM would indeed depend on LTM representations, but all of the heavy lifting would be done by processes outside of LTM itself. The theories of Cowan and Oberauer, which have often been cited as models where STM depends on activated LTM, have additional mechanisms to support binding and serial recall. Indeed, serial recall should probably be seen as the natural task to investigate the relation between verbal STM and LTM. A simpler task where there is no need to represent order might conceivably be performed largely using LTM. Recalling the words “buffalo cowboy cowboy buffalo” in the correct order requires a memory for multiple tokens, but judging that the list contains the word “cowboy” does not. It is also important to bear in mind that factors such as phonological similarity can have different effects on memory for order information and item information. Although phonological similarity generally has a detrimental effect on recall of order, it can have a beneficial effect on item memory (Fallon et al., 1999; Gathercole, Gardiner, & Gregg, 1982; Gupta, Lipinski, & Aktunc, 2005).

Although we can be sure that an activation-based single store model will not work, representations in LTM have an important role to play in assisting STM. One of the big questions here is whether that assistance takes the form of a continuous interaction between LTM and STM during retention, or a reintegration process operating only at recall. Note that if activation from LTM is assumed to improve retention of order, that source of activation needs its own ordering and binding mechanism. An even bigger question though is “What is stored in STM?”. The primary options are copies and pointers. But copies of what, and pointers to what? It is clear that what is stored in the phonological store of the Working Memory model are copies of speech-based representations. However, in the Burgess and Hitch model which is intended

as an implementation of the phonological store, the weights can be thought of as pointers.

Interestingly, the computational properties of this model would remain unchanged if it were assumed that the item representations that were pointed to were in LTM rather than a separate STS. That is, the change would simply be in terms of how the models were described or interpreted, or how they might be implemented in the brain, not in terms of how they worked. As noted earlier though, there are advantages to using pointers within a dedicated STS. Using pointers to local representations can make economical use of storage. It could also minimize the need for information to be transferred between different brain regions. This would be advantageous given that most of the connections within the brain are local. Together, these engineering constraints make a case for expecting the existence of modality specific stores. That is, while it may seem more parsimonious to have a single store, it may be more complicated to build one.

In experimental psychology, much of the progress in understanding STM has followed from the development of computational models. Before the advent of computational theories the Working Memory model had no proper account of how serial recall was actually performed, even though much of the data that motivated the model came from the serial recall task. However, models such as those of Burgess and Hitch (1992), and Page and Norris (1998b) are implementations of the phonological store component of the WM model. Those models can simulate a wide range of data from serial recall. Theories making claims that STM is supported by activated LTM need to rise to the same challenge. This is not simply a matter of requiring them to simulate the data, the models need to be specified in more detail to even begin understand what claims about activated LTM actually mean.

I have concentrated primarily on one possible class of single-store model, those relying solely on activation. Might it be possible to produce more complex single-store models that wouldn't fall foul of the problem of two? The pointer-based models of the form proposed by Ruchkin et al. could represent multiple tokens, but only by adding extra mechanism that serves all of the functions of an STS. However, such a model still faces the seemingly insurmountable challenge of accounting for the neuropsychological data from STM patients. Ruchkin et al. tried to downplay the significance of the neuropsychological data by suggesting that patients with deficits in phonological STM deficits "should be indicative of impairments in establishing long-term memories for those representations" (p. 711). In other words, this is an attempt to turn the argument presented earlier on its head. The problem phonological STM patients have in long-term learning of novel words is not because they have poor STM but because they have trouble establishing phonological representations in LTM. Given that the claim is that STM is supported by activated LTM, it is not clear why a problem in establishing new representations should prevent those representations from being activated—which should be all that's necessary to hold information in STM. According to the Ruchkin et al. view, language comprehension should also depend on activation of phonological representations in STM. However, STM patients have good comprehension abilities, suggesting that they have no problems in activating those representations. Furthermore, the patients' primary STM deficit is in remembering sequences of familiar items such as words or digits. This should not place any demands on establishing new representations

in LTM. Additionally, these patients definitely can establish new representations of verbal material in LTM because they have little trouble with paired-associate learning.

More recently, Surprenant and Neath (2009) also presented a case for a single store model. They mounted a spirited attack on nine lines of argument that have been used to support the distinction between STM and LTM. Number nine concerned data from patients with intact STM but impaired LTM, but they offered no explanation of how data from STM patients might be accounted for in a single store model.⁵

Conclusions

The primary argument advanced here is that any STM system must be able to store complex representational structures that have never been encountered before, whether this be novel sequences of words, or novel combinations of visual features. By definition, such representations cannot already exist in LTM, and therefore cannot be stored simply by activating LTM. The most basic requirement for any memory system is to be able to store multiple tokens of the same type. This is essential for even a simple STM task such as serial recall. If LTM is a store of type representations, STM cannot function simply by activating LTM types. Of course, it is possible to argue that LTM has many representations of a given type, but this would be grossly inefficient as it would imply that all representations in LTM are extensively duplicated. The standard solution is to allow memory to either contain multiple copies of representations, which ensures that only those representations being retained at the moment are duplicated, or to contain pointers to the representations of single types in LTM; the same representation in LTM can be pointed to by multiple pointers in STM. However, simple pointers from STM to LTM are not sufficient to solve the binding problem. There must also be provision for the construction and storage of more complex representations, and this must involve something equivalent to variable binding. Whatever is going on when we store information in STM is far more elaborate than activating LTM, or putting copies in boxes. The important question is not simply whether there are separate short-term and long-term stores, but what is stored in STM. If STM contains pointers, the information capacity of STM will depend on the address space—the wider the range of LTM representations that can be stored in STM (pointed to) the more bits will be required to hold those pointers. This provides an additional motivation for modality-specific stores of the form incorporated into Baddeley and Hitch's working memory model. If some set of computations only need access to a limited range of information that has to be stored temporarily (say phonological information) then it will be more memory-efficient to process that information in a store that is specialized to hold only that kind of information.

The suggestion that STM might be nothing more than activated LTM may have the merit of parsimony, but it simply doesn't work. A simple activation process would be unable to solve the problem of two, or to store novel representations. It follows that any model that places an emphasis on storage by activated LTM must be supplemented by some additional mechanism that can represent multiple tokens serial order. That mechanism must be able to

⁵ See Thorn and Page (2009) for a more detailed rebuttal of the Surprenant and Neath arguments.

perform the variable-binding operation required to construct novel representations. That additional mechanism would then amount to what has conventionally been thought of as a short-term store. In fact, the resulting model would look very much like existing computational models of STM. Some may still prefer to describe this by saying that STM is activated LTM. If they make it clear that there must be some additional mechanism, and explain how that mechanism operates, at least we'll know what they mean.

References

- Aben, B., Stapert, S., & Blokland, A. (2012). About the distinction between working memory and short-term memory. *Frontiers in Psychology, 3*, 301. <http://dx.doi.org/10.3389/fpsyg.2012.00301>
- Acheson, D. J., Hamidi, M., Binder, J. R., & Postle, B. R. (2011). A common neural substrate for language production and verbal working memory. *Journal of Cognitive Neuroscience, 23*, 1358–1367. <http://dx.doi.org/10.1162/jocn.2010.21519>
- Acheson, D. J., MacDonald, M. C., & Postle, B. R. (2011). The effect of concurrent semantic categorization on delayed serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 44–59. <http://dx.doi.org/10.1037/a0021205>
- Anderson, J. R. (1983). *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.
- Atkinson, R. C., & Shiffrin, R. M. (1966). *Some two-process models for memory*. Tech. Rep. No. 107. Institute for Mathematical Studies in the Social Sciences, Stanford.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89–195). Amsterdam, the Netherlands: Elsevier. [http://dx.doi.org/10.1016/S0079-7421\(08\)60422-3](http://dx.doi.org/10.1016/S0079-7421(08)60422-3)
- Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American, 225*, 82–90. <http://dx.doi.org/10.1038/scientificamerican0871-82>
- Baddeley, A. D. (1964). Immediate memory and the “perception” of letter sequences. *The Quarterly Journal of Experimental Psychology, 16*, 364–367. <http://dx.doi.org/10.1080/17470216408416394>
- Baddeley, A. D. (1966a). The influence of acoustic and semantic similarity on long-term memory for word sequences. *The Quarterly Journal of Experimental Psychology, 18*, 302–309. <http://dx.doi.org/10.1080/14640746608400047>
- Baddeley, A. D. (1966b). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *The Quarterly Journal of Experimental Psychology, 18*, 362–365. <http://dx.doi.org/10.1080/14640746608400055>
- Baddeley, A. D. (1986). *Working memory*. Oxford, UK: Oxford University Press.
- Baddeley, A. (1992). Working memory. *Science, 255*, 556–559. <http://dx.doi.org/10.1126/science.1736359>
- Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*, 417–423. [http://dx.doi.org/10.1016/S1364-6613\(00\)01538-2](http://dx.doi.org/10.1016/S1364-6613(00)01538-2)
- Baddeley, A. D. (2003). New data: Old pitfalls. Commentary on Ruchkin, Grafman, Cameron & Berndt. *Behavioral and Brain Sciences, 26*, 728–729. <http://dx.doi.org/10.1017/S0140525X03230168>
- Baddeley, A. D., & Ecob, J. R. (1970). Simultaneous acoustic and semantic coding in short-term memory. *Nature, 227*, 288–289. <http://dx.doi.org/10.1038/227288a0>
- Baddeley, A. D., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158–173. <http://dx.doi.org/10.1037/0033-295X.105.1.158>
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation*. New York, NY: Academic Press. [http://dx.doi.org/10.1016/S0079-7421\(08\)60452-1](http://dx.doi.org/10.1016/S0079-7421(08)60452-1)
- Baddeley, A. D., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology, 36A*, 233–252. <http://dx.doi.org/10.1080/14640748408402157>
- Baddeley, A. D., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory and Language, 27*, 586–595. [http://dx.doi.org/10.1016/0749-596X\(88\)90028-9](http://dx.doi.org/10.1016/0749-596X(88)90028-9)
- Baddeley, A. D., & Warrington, E. K. (1970). Amnesia and the distinction between long- and short-term memory. *Journal of Verbal Learning & Verbal Behavior, 9*, 176–189. [http://dx.doi.org/10.1016/S0022-5371\(70\)80048-2](http://dx.doi.org/10.1016/S0022-5371(70)80048-2)
- Basso, A., Spinnler, H., Vallar, G., & Zanobio, M. E. (1982). Left hemisphere damage and selective impairment of auditory verbal short-term memory. A case study. *Neuropsychologia, 20*, 263–274. [http://dx.doi.org/10.1016/0028-3932\(82\)90101-4](http://dx.doi.org/10.1016/0028-3932(82)90101-4)
- Bettencourt, K. C., & Xu, Y. (2016). Decoding the content of visual short-term memory under distraction in occipital and parietal areas. *Nature Neuroscience, 19*, 150–157. <http://dx.doi.org/10.1038/nn.4174>
- Bormann, T., Seyboth, M., Umarova, R., & Weiller, C. (2015). “I know your name, but not your number”—Patients with verbal short-term memory deficits are impaired in learning sequences of digits. *Neuropsychologia, 72*, 80–86. <http://dx.doi.org/10.1016/j.neuropsychologia.2015.03.027>
- Botvinick, M. M. (2005). Effects of domain-specific knowledge on memory for serial order. *Cognition, 97*, 135–151. <http://dx.doi.org/10.1016/j.cognition.2004.09.007>
- Botvinick, M., & Bylsma, L. M. (2005). Regularization in short-term memory for serial order. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 351–358. <http://dx.doi.org/10.1037/0278-7393.31.2.351>
- Botvinick, M. M., & Plaut, D. C. (2006). Short-term memory for serial order: A recurrent neural network model. *Psychological Review, 113*, 201–233. <http://dx.doi.org/10.1037/0033-295X.113.2.201>
- Bourassa, D. C., & Besner, D. (1994). Beyond the articulatory loop: A semantic contribution to serial order recall of subspan lists. *Psychonomic Bulletin & Review, 1*, 122–125. <http://dx.doi.org/10.3758/BF03200768>
- Bowman, H., & Wyble, B. (2007). The simultaneous type, serial token model of temporal attention and working memory. *Psychological Review, 114*, 38–70. <http://dx.doi.org/10.1037/0033-295X.114.1.38>
- Brener, R. (1940). An experimental investigation of memory span. *Journal of Experimental Psychology, 26*, 467–482. <http://dx.doi.org/10.1037/h0061096>
- Brown, G. D., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review, 114*, 539–576. <http://dx.doi.org/10.1037/0033-295X.114.3.539>
- Brown, G. D., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review, 107*, 127–181. <http://dx.doi.org/10.1037/0033-295X.107.1.127>
- Bullmore, E., & Sporns, O. (2012). The economy of brain network organization. *Nature Reviews Neuroscience, 13*, 336–349.
- Burgess, N., & Hitch, G. J. (1992). Toward a network model of the articulatory loop. *Journal of Memory and Language, 31*, 429–460. [http://dx.doi.org/10.1016/0749-596X\(92\)90022-P](http://dx.doi.org/10.1016/0749-596X(92)90022-P)
- Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106*, 551–581. <http://dx.doi.org/10.1037/0033-295X.106.3.551>
- Burgess, N., & Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language, 55*, 627–652. <http://dx.doi.org/10.1016/j.jml.2006.08.005>
- Cameron, K. A., Haarmann, H. J., Grafman, J., & Ruchkin, D. S. (2005). Long-term memory is the representational basis for semantic verbal

- short-term memory. *Psychophysiology*, 42, 643–653. <http://dx.doi.org/10.1111/j.1469-8986.2005.00357.x>
- Cantor, J., & Engle, R. W. (1993). Working-memory capacity as long-term memory activation: An individual-differences approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 1101–1114. <http://dx.doi.org/10.1037/0278-7393.19.5.1101>
- Cave, C. B., & Squire, L. R. (1992). Intact verbal and nonverbal short-term memory following damage to the human hippocampus. *Hippocampus*, 2, 151–163. <http://dx.doi.org/10.1002/hipo.450020207>
- Christophel, T. B., Hebart, M. N., & Haynes, J.-D. (2012). Decoding the contents of visual short-term memory from human visual and parietal cortex. *The Journal of Neuroscience*, 32, 12983–12989. <http://dx.doi.org/10.1523/JNEUROSCI.0184-12.2012>
- Conrad, R. (1965). Order error in immediate recall of sequences. *Journal of Verbal Learning and Verbal Behavior*, 4, 161–169. [http://dx.doi.org/10.1016/S0022-5371\(65\)80015-9](http://dx.doi.org/10.1016/S0022-5371(65)80015-9)
- Corsi, P. M. (1973). *Human memory and the medial temporal region of the brain*. ProQuest Information & Learning.
- Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, 104, 163–191. <http://dx.doi.org/10.1037/0033-2909.104.2.163>
- Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge, UK: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139174909.006>
- Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research*, 169, 323–338. [http://dx.doi.org/10.1016/S0079-6123\(07\)00020-9](http://dx.doi.org/10.1016/S0079-6123(07)00020-9)
- Cowan, N., & Chen, Z. (2008). How chunks form in long-term memory and affect short-term memory limits. In M. Page & A. Thorn (Eds.), *Interactions between short-term and long-term memory in the verbal domain* (pp. 86–107). Hove, UK: Psychology Press.
- Crowder, R. G. (1993). Short-term memory: Where do we stand? *Memory & Cognition*, 21, 142–145. <http://dx.doi.org/10.3758/BF03202725>
- Crowder, R. G., & Neath, I. (1991). The microscope metaphor in memory. In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honour of Bennet B. Murdock Jr* (pp. 111–125). Hillsdale, NJ: Erlbaum.
- Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition*, 42(1–3), 287–314. [http://dx.doi.org/10.1016/0010-0277\(92\)90046-K](http://dx.doi.org/10.1016/0010-0277(92)90046-K)
- D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. *Annual Review of Psychology*, 66, 115–142. <http://dx.doi.org/10.1146/annurev-psych-010814-015031>
- Dittmann, J., & Abel, S. (2010). Verbales Arbeitsgedächtnis und verbales Lernen: Wort- und Pseudowortlernen in einem Fall von pathologischer Arbeitsgedächtnisbeeinträchtigung [Verbal working memory and verbal learning: Word- and pseudoword learning in a case of pathological verbal working memory impairment]. *Sprache, Stimme, Gehör*, 34, 1–9. <http://dx.doi.org/10.1055/s-0030-1253361>
- Drachman, D. A., & Arbit, J. (1966). Memory and the hippocampal complex. II. Is memory a multiple process? *Archives of Neurology*, 15, 52–61. <http://dx.doi.org/10.1001/archneur.1966.00470130056005>
- Engle, R. W., Kane, M. J., & Tuoholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). London, England: Cambridge Press. <http://dx.doi.org/10.1017/CBO9781139174909.007>
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102, 211–245. <http://dx.doi.org/10.1037/0033-295X.102.2.211>
- Fallon, A. B., Groves, K., & Tehan, G. (1999). Phonological similarity and trace degradation in the serial recall task: When CAT helps RAT, but not MAN. *International Journal of Psychology*, 34(5–6), 301–307. <http://dx.doi.org/10.1080/002075999399602>
- Farrell, S. (2012). Temporal clustering and sequencing in short-term memory and episodic memory. *Psychological Review*, 119, 223–271. <http://dx.doi.org/10.1037/a0027371>
- Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review*, 9, 59–79. <http://dx.doi.org/10.3758/BF03196257>
- Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28, 3–71. [http://dx.doi.org/10.1016/0010-0277\(88\)90031-5](http://dx.doi.org/10.1016/0010-0277(88)90031-5)
- Freedman, M. L., & Martin, R. C. (2001). Dissociable components of short-term memory and their relation to long-term learning. *Cognitive Neuropsychology*, 18, 193–226. <http://dx.doi.org/10.1080/02643290126002>
- Gathercole, S. E., Gardiner, J. M., & Gregg, V. H. (1982). Modality and phonological similarity effects in serial recall: Does one's own voice play a role? *Memory & Cognition*, 10, 176–180. <http://dx.doi.org/10.3758/BF03209219>
- Gathercole, S. E., Pickering, S. J., Hall, M., & Peaker, S. M. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *Quarterly Journal of Experimental Psychology*, 54, 1–30. <http://dx.doi.org/10.1080/02724980042000002>
- Gregg, V. H., Freedman, C. M., & Smith, D. K. (1989). Word frequency, articulatory suppression and memory span. *British Journal of Psychology*, 80, 363–374. <http://dx.doi.org/10.1111/j.2044-8295.1989.tb02326.x>
- Gruneberg, M. M. (1970). A dichotomous theory of memory - Unproved and unproveable. *Acta Psychologica*, 34, 489–496. [http://dx.doi.org/10.1016/0001-6918\(70\)90042-9](http://dx.doi.org/10.1016/0001-6918(70)90042-9)
- Gupta, P., Lipinski, J., & Aktunc, E. (2005). Reexamining the phonological similarity effect in immediate serial recall: The roles of type of similarity, category cuing, and item recall. *Memory & Cognition*, 33, 1001–1016. <http://dx.doi.org/10.3758/BF03193208>
- Hanley, J. R., Young, A. W., & Pearson, N. A. (1991). Impairment of the visuo-spatial sketch pad. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, 43, 101–125. <http://dx.doi.org/10.1080/14640749108401001>
- Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, 458, 632–635. <http://dx.doi.org/10.1038/nature07832>
- Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye (Ed.), *Brain mechanisms and learning* (pp. 37–46). Oxford, UK: Blackwell.
- Henson, R. N. (1998). Short-term memory for serial order: The Start-End Model. *Cognitive Psychology*, 36, 73–137. <http://dx.doi.org/10.1006/cogp.1998.0685>
- Hulme, C., Maughan, S., & Brown, G. D. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, 30, 685–701. [http://dx.doi.org/10.1016/0749-596X\(91\)90032-F](http://dx.doi.org/10.1016/0749-596X(91)90032-F)
- Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D., Martin, M., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 1217–1232. <http://dx.doi.org/10.1037/0278-7393.23.5.1217>
- Huttenlocher, J., & Newcombe, N. (1976). Semantic effects on ordered recall. *Journal of Verbal Learning & Verbal Behavior*, 15, 387–399. [http://dx.doi.org/10.1016/S0022-5371\(76\)90034-7](http://dx.doi.org/10.1016/S0022-5371(76)90034-7)

- Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, 92(1–2), 101–144. <http://dx.doi.org/10.1016/j.cognition.2002.06.001>
- Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford, UK: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780198270126.001.0001>
- James, W. (1890). *The principles of psychology* (2 Vol.). New York, NY: Henry Holt (Reprinted Bristol, England: Thoemmes Press, 1999).
- Jefferies, E., Frankish, C., & Lambon-Ralph, M. A. (2006a). Lexical and semantic binding in verbal short-term memory. *Journal of Memory and Language*, 54, 81–98. <http://dx.doi.org/10.1016/j.jml.2005.08.001>
- Jefferies, E., Frankish, C., & Lambon Ralph, M. A. (2006b). Lexical and semantic influences on item and order memory in immediate serial recognition: Evidence from a novel task. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 59, 949–964. <http://dx.doi.org/10.1080/02724980543000141>
- Jefferies, E., Frankish, C., & Noble, K. (2009). Lexical coherence in short-term memory: Strategic reconstruction or “semantic glue”? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 62, 1967–1982. <http://dx.doi.org/10.1080/17470210802697672>
- Jenison, A., & Squire, L. R. (2012). Working memory, long-term memory, and medial temporal lobe function. *Learning & Memory*, 19, 15–25. <http://dx.doi.org/10.1101/lm.024018.111>
- Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. S. (2008). The mind and brain of short-term memory. *Annual Review of Psychology*, 59, 193–224. <http://dx.doi.org/10.1146/annurev.psych.59.103006.093615>
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99, 122–149. <http://dx.doi.org/10.1037/0033-295X.99.1.122>
- Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & R. Davies (Eds.), *Varieties of attention* (pp. 29–61). New York, NY: Academic Press.
- Kalm, K., Davis, M. H., & Norris, D. (2013). Individual sequence representations in the medial temporal lobe. *Journal of Cognitive Neuroscience*, 25, 1111–1121. http://dx.doi.org/10.1162/jocn_a_00378
- Kanwisher, N. (1987). Repetition blindness - Type recognition without token individuation. *Cognition*, 27, 117–143. [http://dx.doi.org/10.1016/0010-0277\(87\)90016-3](http://dx.doi.org/10.1016/0010-0277(87)90016-3)
- Kanwisher, N. (1991). Repetition blindness and illusory conjunctions: Errors in binding visual types with visual tokens. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 404–421. <http://dx.doi.org/10.1037/0096-1523.17.2.404>
- LaRocque, J. J., Eichenbaum, A. S., Starrett, M. J., Rose, N. S., Emrich, S. M., & Postle, B. R. (2015). The short- and long-term fates of memory items retained outside the focus of attention. *Memory & Cognition*, 43, 453–468. <http://dx.doi.org/10.3758/s13421-014-0486-y>
- LaRocque, J. J., & Lewis-Peacock, J. A. (2013). Decoding attended information in short-term memory: An EEG study. *Journal of Cognitive Neuroscience*, 25, 127–142. <http://dx.doi.org/10.1162/jocn>
- LaRocque, J. J., Lewis-Peacock, J. A., & Postle, B. R. (2014). Multiple neural states of representation in short-term memory? It's a matter of attention. *Frontiers in Human Neuroscience*, 8, 5. <http://dx.doi.org/10.3389/fnhum.2014.00005>
- LaRocque, J. J., Riggall, A. C., Emrich, S. M., & Postle, B. R. (2016). Within-category decoding of information in different attentional states in short-term memory. *Cerebral Cortex*. Advance online publication. <http://dx.doi.org/10.1093/cercor/bhw283>
- Lee, S.-H., & Baker, C. I. (2016). Multi-voxel decoding and the topography of maintained information during visual working memory. *Frontiers in Systems Neuroscience*, 10, 2. <http://dx.doi.org/10.3389/fnsys.2016.00002>
- Levelt, W. J. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition*, 42(1–3), 1–22. [http://dx.doi.org/10.1016/0010-0277\(92\)90038-J](http://dx.doi.org/10.1016/0010-0277(92)90038-J)
- Lewandowsky, S., & Farrell, S. (2000). A reintegration account of the effects of speech rate, lexicality, and word frequency in immediate serial recall. *Psychological Research*, 63, 163–173. <http://dx.doi.org/10.1007/PL00008175>
- Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of Cognitive Neuroscience*, 24, 61–79. http://dx.doi.org/10.1162/jocn_a_00140
- Lewis-Peacock, J. A., & Postle, B. R. (2008). Temporary activation of long-term memory supports working memory. *The Journal of Neuroscience*, 28, 8765–8771. <http://dx.doi.org/10.1523/JNEUROSCI.1953-08.2008>
- Logie, R. H. (2003). Spatial and visual working memory: A mental workspace. *Psychology of Learning and Motivation*, 42, 37–78. [http://dx.doi.org/10.1016/S0079-7421\(03\)01002-8](http://dx.doi.org/10.1016/S0079-7421(03)01002-8)
- Logie, R. H., & Della Sala, S. (2003). Working memory as a mental workspace: Why activated long-term memory is not enough. *Behavioral and Brain Sciences*, 26, 745–746. <http://dx.doi.org/10.1017/S0140525X03400162>
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279–281. <http://dx.doi.org/10.1038/36846>
- Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, 17, 391–400. <http://dx.doi.org/10.1016/j.tics.2013.06.006>
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. Cambridge, MA: MIT Press.
- Martin, R. C., & Romani, C. (1994). Verbal working memory and sentence comprehension: A multiple-components view. *Neuropsychology*, 8, 506–523. <http://dx.doi.org/10.1037/0894-4105.8.4.506>
- Martin, R. C., Shelton, J. R., & Yaffee, L. S. (1994). Language processing and working memory: Neuropsychological evidence for separate phonological and semantic capacities. *Journal of Memory and Language*, 33, 83–111. <http://dx.doi.org/10.1006/jmla.1994.1005>
- Mathy, F., & Feldman, J. (2012). What's magic about magic numbers? Chunking and data compression in short-term memory. *Cognition*, 122, 346–362. <http://dx.doi.org/10.1016/j.cognition.2011.11.003>
- McElree, B. (2006). Accessing recent events. *Psychology of Learning and Motivation*, 46, 155–200. [http://dx.doi.org/10.1016/S0079-7421\(06\)46005-9](http://dx.doi.org/10.1016/S0079-7421(06)46005-9)
- Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 2, 1–21. [http://dx.doi.org/10.1016/S0022-5371\(63\)80063-8](http://dx.doi.org/10.1016/S0022-5371(63)80063-8)
- Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology*, 78, 679–684. <http://dx.doi.org/10.1037/h0026641>
- Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology*, 53, 53–81. <http://dx.doi.org/10.1146/annurev.psych.53.100901.135131>
- Nee, D. E., & Jonides, J. (2013). Trisecting representational states in short-term memory. *Frontiers in Human Neuroscience*, 7, 796. <http://dx.doi.org/10.3389/fnhum.2013.00796>
- Norman, D. (1968). Toward a theory of memory and attention. *Psychological Review*, 75, 522–536. <http://dx.doi.org/10.1037/h0026699>
- Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 411–421. <http://dx.doi.org/10.1037/0278-7393.28.3.411>

- Oberauer, K. (2009). Design for a working memory. *Psychology of Learning and Motivation*, 51, 45–100. [http://dx.doi.org/10.1016/S0079-7421\(09\)51002-X](http://dx.doi.org/10.1016/S0079-7421(09)51002-X)
- Öztekin, I., Davachi, L., & McElree, B. (2010). Are representations in working memory distinct from representations in long-term memory? Neural evidence in support of a single store. *Psychological Science*, 21, 1123–1133. <http://dx.doi.org/10.1177/0956797610376651>
- Page, M., Madge, A., Cumming, N., & Norris, D. (2007). Speech errors and the phonological similarity effect in short-term memory: Evidence suggesting a common locus ☆. *Journal of Memory and Language*, 56, 49–64. <http://dx.doi.org/10.1016/j.jml.2006.09.002>
- Page, M., & Norris, D. (1998a). Modeling immediate serial recall with a localist implementation of the primacy model. In J. Grainger & A. M. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 227–255). Mahwah, NJ: Erlbaum.
- Page, M. P., & Norris, D. (1998b). The primacy model: A new model of immediate serial recall. *Psychological Review*, 105, 761–781. <http://dx.doi.org/10.1037/0033-295X.105.4.761-781>
- Page, M. P., & Norris, D. (2009). A model linking immediate serial recall, the Hebb repetition effect and the learning of phonological word forms. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*, 364, 3737–3753. <http://dx.doi.org/10.1098/rstb.2009.0173>
- Patterson, K., Graham, N., & Hodges, J. R. (1994). The impact of semantic memory loss on phonological representations. *Journal of Cognitive Neuroscience*, 6, 57–69. <http://dx.doi.org/10.1162/jocn.1994.6.1.57>
- Peterson, L. R., & Johnson, S. T. (1971). Some effects of minimizing articulation on short-term retention. *Journal of Verbal Learning & Verbal Behavior*, 10, 346–354. [http://dx.doi.org/10.1016/S0022-5371\(71\)80033-6](http://dx.doi.org/10.1016/S0022-5371(71)80033-6)
- Phillips, W. A. (2003). The short-term dynamics within a network of connections is creative. Commentary on Ruchkin, Grafman, Cameron & Berndt. *Behavioral and Brain Sciences*, 26, 752–753. <http://dx.doi.org/10.1017/S0140525X03480163>
- Poirier, M., & Saint-Aubin, J. (1996). Immediate serial recall, word frequency, item identity and item position. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 50, 408–412. <http://dx.doi.org/10.1037/1196-1961.50.4.408>
- Poirier, M., Saint-Aubin, J., Mair, A., Tehan, G., & Tolan, A. (2015). Order recall in verbal short-term memory: The role of semantic networks. *Memory & Cognition*, 43, 489–499. <http://dx.doi.org/10.3758/s13421-014-0470-6>
- Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, 139, 23–38. <http://dx.doi.org/10.1016/j.neuroscience.2005.06.005>
- Potter, M. C. (1993). Very short-term conceptual memory. *Memory & Cognition*, 21, 156–161. <http://dx.doi.org/10.3758/BF03202727>
- Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, 32, 65–97. [http://dx.doi.org/10.1016/0010-0277\(89\)90014-0](http://dx.doi.org/10.1016/0010-0277(89)90014-0)
- Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, 80(1–2), 127–158. [http://dx.doi.org/10.1016/S0010-0277\(00\)00156-6](http://dx.doi.org/10.1016/S0010-0277(00)00156-6)
- Ranganath, C., & Blumenfeld, R. S. (2005). Doubts about double dissociations between short- and long-term memory. *Trends in Cognitive Sciences*, 9, 374–380. <http://dx.doi.org/10.1016/j.tics.2005.06.009>
- Ruchkin, D. S., Grafman, J., Cameron, K., & Berndt, R. S. (2003). Working memory retention systems: A state of activated long-term memory. *Behavioral and Brain Sciences*, 26, 709–728. <http://dx.doi.org/10.1017/S0140525X03000165>
- Saffran, E. M., & Marin, O. S. (1975). Immediate memory for word lists and sentences in a patient with deficient auditory short-term memory. *Brain and Language*, 2, 420–433. [http://dx.doi.org/10.1016/S0093-934X\(75\)80081-2](http://dx.doi.org/10.1016/S0093-934X(75)80081-2)
- Saint-Aubin, J., & Poirier, M. (1999a). The influence of long-term memory factors on immediate serial recall: An item and order analysis. *International Journal of Psychology*, 34, 347–352. <http://dx.doi.org/10.1080/002075999399675>
- Saint-Aubin, J., & Poirier, M. (1999b). Semantic similarity and immediate serial recall: Is there a detrimental effect on order information? *The Quarterly Journal of Experimental Psychology: Section A*, 52, 367–394. <http://dx.doi.org/10.1080/713755814>
- Saint-Aubin, J., & Poirier, M. (2000). Immediate serial recall of words and nonwords: Tests of the retrieval-based hypothesis. *Psychonomic Bulletin & Review*, 7, 332–340. <http://dx.doi.org/10.3758/BF03212990>
- Savill, N., Metcalfe, T., Ellis, A. W., & Jefferies, E. (2015). Semantic categorisation of a word supports its phonological integrity in verbal short-term memory. *Journal of Memory and Language*, 84, 128–138. <http://dx.doi.org/10.1016/j.jml.2015.06.003>
- Schweickert, R. (1993). A multinomial processing tree model for degradation and reintegration in immediate recall. *Memory & Cognition*, 21, 168–175. <http://dx.doi.org/10.3758/BF03202729>
- Schweickert, R., Chen, S., & Poirier, M. (1999). Redintegration and the useful lifetime of the verbal memory representation. *International Journal of Psychology*, 34, 447–453. <http://dx.doi.org/10.1080/002075999399800>
- Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, & Psychiatry*, 20, 11–21. <http://dx.doi.org/10.1136/jnnp.20.1.11>
- Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science*, 20, 207–214. <http://dx.doi.org/10.1111/j.1467-9280.2009.02276.x>
- Shallice, T., & Vallar, G. (1990). The impairment of auditory-verbal short-term storage. In T. Shallice & G. Vallar (Eds.), *Neuropsychological impairments of short-term memory* (pp. 11–53). Cambridge, UK: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511665547.003>
- Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *The Quarterly Journal of Experimental Psychology*, 22, 261–273. <http://dx.doi.org/10.1080/00335557043000203>
- Shannon, C. E. (1951). Prediction and entropy of printed English. *The Bell System Technical Journal*, 30, 50–64. <http://dx.doi.org/10.1002/j.1538-7305.1951.tb01366.x>
- Shiffrin, R. M. (1975). Capacity limitations in information processing, attention and memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes: Memory processes* (Vol. 4, pp. 177–236). Hillsdale, NJ: Erlbaum.
- Shivde, G., & Anderson, M. C. (2011). On the existence of semantic working memory: Evidence for direct semantic maintenance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 1342–1370. <http://dx.doi.org/10.1037/a0024832>
- Shivde, G., & Thompson-Schill, S. L. (2004). Dissociating semantic and phonological maintenance using fMRI. *Cognitive, Affective & Behavioral Neuroscience*, 4, 10–19. <http://dx.doi.org/10.3758/CABN.4.1.10>
- Sprague, T. C., Ester, E. F., & Serences, J. T. (2014). Reconstructions of information in visual spatial working memory degrade with memory load. *Current Biology*, 24, 2174–2180. <http://dx.doi.org/10.1016/j.cub.2014.07.066>
- Sprague, T. C., Ester, E. F., & Serences, J. T. (2016). Restoring latent visual working memory representations in human cortex. *Neuron*, 91, 694–707. <http://dx.doi.org/10.1016/j.neuron.2016.07.006>
- Stilp, C. E. (2011). The redundancy of phonemes in sentential context. *The Journal of the Acoustical Society of America*, 130, EL323–EL328. <http://dx.doi.org/10.1121/1.3645966>
- Stuart, G., & Hulme, C. (2000). The effects of word co-occurrence on short-term memory: Associative links in long-term memory affect short-

- term memory performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 796–802. <http://dx.doi.org/10.1037/0278-7393.26.3.796>
- Surprenant, A. M., & Neath, I. (2009). The nine lives of short-term memory. In A. Thorn & M. Page (Eds.), *Interactions between short-term and long-term memory in the verbal domain* (pp. 16–43). Hove, UK: Psychology Press.
- Talmi, D., Grady, C. L., Goshen-Gottstein, Y., & Moscovitch, M. (2005). Neuroimaging the serial position curve. A test of single-store versus dual-store models. *Psychological Science*, 16, 716–723. <http://dx.doi.org/10.1111/j.1467-9280.2005.01601.x>
- Thorn, A., & Page, M. (2009). Current issues in understanding interactions between short-term and long-term memory. In A. Thorn & M. Page (Eds.), *Interactions between short-term and long-term memory in the verbal domain* (pp. 1–15). Hove, UK: Psychology Press.
- Trojano, L., & Grossi, D. (1995). Phonological and lexical coding in verbal short-term memory and learning. *Brain and Language*, 51, 336–354. <http://dx.doi.org/10.1006/brln.1995.1064>
- Trojano, L., Stanzione, M., & Grossi, D. (1992). Short-term memory and verbal learning with auditory phonological coding defect: A neuropsychological case study. *Brain and Cognition*, 18, 12–33. [http://dx.doi.org/10.1016/0278-2626\(92\)90108-X](http://dx.doi.org/10.1016/0278-2626(92)90108-X)
- Vallar, G., & Baddeley, A. D. (1984). Fractionation of working memory: Neuropsychological evidence for a phonological short-term store. *Journal of Verbal Learning & Verbal Behavior*, 23, 151–161. [http://dx.doi.org/10.1016/S0022-5371\(84\)90104-X](http://dx.doi.org/10.1016/S0022-5371(84)90104-X)
- Vallar, G., Di Betta, A. M., & Silveri, M. C. (1997). The phonological short-term store-rehearsal system: Patterns of impairment and neural correlates. *Neuropsychologia*, 35, 795–812. [http://dx.doi.org/10.1016/S0028-3932\(96\)00127-3](http://dx.doi.org/10.1016/S0028-3932(96)00127-3)
- Vallar, G., & Papagno, C. (1995). Neuropsychological impairments of short-term memory. In A. D. Baddeley, B. A. Wilson, & F. N. Wats (Eds.), *Handbook of memory disorders*. New York, NY: Wiley.
- Vallar, G., Papagno, C., & Baddeley, A. D. (1991). Long-term recency effects and phonological short-term memory. A neuropsychological case study. *Cortex*, 27, 323–326. [http://dx.doi.org/10.1016/S0010-9452\(13\)80137-2](http://dx.doi.org/10.1016/S0010-9452(13)80137-2)
- van der Velde, F., & de Kamps, M. (2006). Neural blackboard architectures of combinatorial structures in cognition. *Behavioral and Brain Sciences*, 29, 37–70. <http://dx.doi.org/10.1017/S0140525X06009022>
- Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 1256–1271. <http://dx.doi.org/10.1037/0278-7393.25.5.1256>
- Warrington, E. K., Logue, V., & Pratt, R. T. (1971). The anatomical localisation of selective impairment of auditory verbal short-term memory. *Neuropsychologia*, 9, 377–387. [http://dx.doi.org/10.1016/0028-3932\(71\)90002-9](http://dx.doi.org/10.1016/0028-3932(71)90002-9)
- Warrington, E. K., & Shallice, T. (1969). The selective impairment of auditory verbal short-term memory. *Brain: A Journal of Neurology*, 92, 885–896. <http://dx.doi.org/10.1093/brain/92.4.885>
- Watkins, O. C., & Watkins, J. M. (1977). Serial recall and the modality effect: Effects of word frequency. *Journal of Experimental Psychology: Human Learning and Memory*, 3, 712–718. <http://dx.doi.org/10.1037/0278-7393.3.6.712>
- Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review*, 72, 89–104. <http://dx.doi.org/10.1037/h0021797>
- Wickelgren, W. A. (1973). The long and the short of memory. *Psychological Bulletin*, 80, 425–438. <http://dx.doi.org/10.1037/h0035255>
- Wilson, B., & Baddeley, A. (1988). Semantic, episodic, and autobiographical memory in a postmeningitic amnesic patient. *Brain and Cognition*, 8, 31–46. [http://dx.doi.org/10.1016/0278-2626\(88\)90037-1](http://dx.doi.org/10.1016/0278-2626(88)90037-1)

Received March 31, 2016

Revision received December 8, 2016

Accepted April 16, 2017 ■

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <https://my.apa.org/portal/alerts/> and you will be notified by e-mail when issues of interest to you become available!