15

Multi-Store versus Dynamic Models of Temporary Storage in Memory

Randall W. Engle and Natalie Oransky

The concept of short-term memory as a distinct type of memory has now become part of our cultural common knowledge. Evidence the Tom Hanks character Mr. Short-Term Memory on the NBC television show *Saturday Night Live.* Mr. STM has good retention of very recent events, but that information is quickly lost and he shows no evidence of long-term retention. However, there has been strong debate throughout the entire era of modern memory research about whether memory for recent events and long-term retention obey different principles. Proponents of multiple memory stores suggest that different memory structures yield distinct memory traces. It seems intuitive that we have several types of memory (e.g., memories that are short lived and others that last for a very long time), but some theorists argue that it is not necessary to posit separate structures to explain different memory traces. According to this unitary approach to memory, differences between memories occur not because they are stored in separate systems, but because of the different processes and modes of representation used to perceive or think about events when they occur. Although the issue of whether or not memory is composed of stages or components is prevalent throughout the memory literature (e.g., implicit and explicit memory; see Roediger, 1990, and Schacter, 1987); this chapter will focus on the contribution of these opposing views to the development of the construct of short-term memory and its younger relative, *working memory.*

In this chapter, we will present a rather cursory history of the ideas about short-term memory along with some of the research findings that were presented to support or refute those ideas. The literature on short-term memory was an important part of the experimental psychology of human memory in the 1960s and early '70s. Early research focused on the nature of forgetting and whether forgetting from short-term memory is a result of decay, displacement, or some other mechanism different from those thought to mediate forgetting from long-term memory. One of the stronger proposals about short-term memory from the earliest theories (e.g., Hebb, 1949) was that the strength of a trace in long-term memory is a function of how long the information is retained in short-term memory. We will see in a discussion of transfer of information to long-term memory that that proposal is quite wrong. Early research also suggested that short-term memory is characterized by a speech-based code, whereas long-term memory is characterized by meaning-based codes. That idea also is likely wrong. These concerns gave rise to a disenchantment with the concept of short-term memory. But the seminal paper by Baddeley and Hitch (1974), along with new ideas in neo-Piagetian approaches to developmental psychology (Pascal-Leone, 1970) and brain-based approaches to memory, led to a renascence that makes working memory one of the core topics in cognitive and developmental psychology today.

Multi-Store Model

The information-processing model proposed by Atkinson and Shiffrin (1968) represents a prototypical multi-store approach to memory. This model distinguished between sensory, short-term, and long-term stores, and it suggested that information is processed through these structurally and functionally independent stages. At the completion of each stage, products are copied to the next stage for further processing. Atkinson and Shiffrin's model combined the attentional system and the temporary storage component of Broadbent's (1958) seminal model into one limited capacity, short-term storage system. Their *short-term store* (STS) has a limited number of slots for holding information,[1] and this information decays rapidly if it is not maintained by control processes such as rehearsal. Support for multi-store models came from data showing differences between the characteristics of the proposed memory systems. For example, Atkinson and Shiffrin characterized STS as having a limited capacity, using primarily verbal codes, and losing information due to decay. In contrast, *long-term store* (LTS) was characterized as having a large capacity, primarily semantic coding, and losing information via interference.

Forgetting

It is difficult to understand the impetus for much of the early work on STS unless you understand that many of the putative characteristics of short-term memory arise from a theory about the physiology of the memory trace put forth by Hebb (1949). Hebb proposed that if two events occur in close proximity to one another, the neural circuits corresponding to those events are active at the same time, and new connections, called *reverberatory traces,* are formed between those two circuits. If given enough uninterrupted time, this new connection will consolidate and become a relatively permanent structural trace. Other events can, however, prevent the consolidation of the new trace so that the connection is lost, or more appropriately, never formed. Hebb envisioned that reverberatory traces and consolidated structural traces are qualitatively different states. The reverberatory trace is time limited and capable of *displacement* by new events. Neither of these traits are true of the long-term structural traces. There was relatively little evidence to support Hebb's ideas at the time and his book was rarely cited in the short-term memory literature that followed, but the influence of those ideas is obvious.[2]

Following Hebb's theory, it is not surprising that the dominant issue in early research on the STS focused on forgetting. Proponents of multi-store approaches demonstrated differences in the mechanisms of information loss from STS and LTS. They suggested that information loss from STS results from one of two types of limitations: temporal persistence and capacity limitations. Some theorists argued that there are limits in *how long* information could be held in STS Brown, 1958; Conrad, 1957; Hebb, 1949; Peterson & Peterson, 1959). That is, if information is not rehearsed, it decays, or fades. Others suggested that there is a limit in the *number* of items that can be held in STS at a given moment and that information is lost because new information replaces or bumps old information out of STS (e.g., James, 1890; Miller, 1956; Waugh & Norman, 1965). Both of these mechanisms of forgetting, time limits and capacity limits, were distinguished from mechanisms proposed to account for long-term forgetting. More specifically, most researchers agreed that information in LTS is long lasting and is lost temporarily through proactive or retroactive interference affecting retrieval rather than through decay or replacement (McGeoch, 1932; Melton & Irwin, 1940; Postman, Stark, & Fraser, 1968).

Brown (1958) and Peterson and Peterson (1959) argued for decay from STS based on the findings that when subjects are prevented from rehearsing, recall quickly declines over the delay. In the Peterson and Peterson experiments, subjects saw trigrams and tried to recall them after a delay period filled with a rehearsal-preventative task requiring backward counting from a number. Recall accuracy of the trigrams dropped dramatically over filled delays of 18 seconds, suggesting that without rehearsal, information in STS is short lived and relatively transient and cannot be transferred to LTS. The rehearsal-preventative task involved numbers, so according to traditional interference theory (McGeoch, 1932), there should not have been material-specific interference for the letters or words in such a task. Further, Peterson and Peterson argued that proactive interference across trials did not occur in their short-term memory task. That is, early trials did not block the retrieval of items on later trials. The rate of forgetting was the same for the first block of 10 trials as for those blocks tested later. However, Keppel and Underwood (1962) showed that the build-up of proactive interference in the Peterson and Peterson task was rapid and quickly reached asymptote within the first 2 to 3 trials of the experiment. Peterson and Peterson had masked this interference by looking only at blocks of 10 trials. These and other results showing interference as a factor in forgetting in short-term memory tasks (e.g., Hebb, 1961; Murdock, 1961) led Melton (1963) to suggest that all memory traces, whether formed 5 seconds ago or 5 years ago, share the same characteristics: namely, they are stable, permanent traces that are susceptible to interference from other traces during retrieval. Speaking to Hebb's reverberatory trace idea, Melton argued for "the fixation of a structural trace by a single repetition of an event without the benefit of autonomous consolidation processes" (p. 19).

Waugh and Norman (1965) used a digit probe task to examine both decay and interference as factors in the loss of information from STS. Subjects heard digits presented at either a one-per-second or a four-per-second rate. At the end of a series of digits, subjects received a probe digit and recalled the digit that had occurred just prior to the probe in the digit series. The number of items intervening between the probe digit and the end of the list was manipulated. Waugh and Norman reasoned that if decay causes forgetting, then recall in the one-item-per-second condition should be worse than recall in the four-items-per-second condition because more time will have elapsed between the presentation of the target and recall. On the other hand, if forgetting is the result of capacity limits, recall of the target digit should be a function of the number of intervening items regardless of presentation rate. Waugh and Norman showed that recall is affected by the number of intervening items and not presentation rate, suggesting that displacement is the critical factor in STS forgetting. Waugh and Norman proposed that STS holds a limited amount of information and once the limit is reached, new information displaces old information, causing old information to be permanently lost. In addition, Waugh and Norman showed that when reinterpreted within their framework, data from other short-term memory tasks (e.g., immediate free recall) support a capacity/displacement interpretation of STS forgetting. This view is consistent with Miller's (1956) conception of STS as containing a fixed number of 7 ± 2 slots or bins for holding information. Thus, Waugh and Norman argued that memory traces in STS are transient because of capacity limitations. However, these traces can be maintained indefinitely in STS and copied to LTS through rehearsal.

Perhaps the most careful analysis of this issue was in two papers by Reitman (1971, 1974). She painstakingly prevented rehearsal over the 15-second filled delay, avoided a ceiling effect for the initial trials she used a task similar to that used by Brown (1958) and Peterson and Peterson (1959) in which each trial presented items for recall followed by a delay before recall,[3] and manipulated the nature of the rehearsal-preventative task, with one task being a tone detection task and the other being a syllable detection task. In the nonverbal but attention-demanding tone detection task, there was a 12% decline in recall, which Reitman attributed to decay. However, the syllable detection task led to a 56% loss in recall. She argued that the predominant cause of forgetting from short-term memory is interference through displacement but that there is evidence also for decay. Thus, there is evidence favoring both time limits and capacity limits as mechanisms for forgetting in STS.

More recent findings also suggest a role for a time-based loss of information. Baddeley, Thomson, and Buchanan (1975) showed that the number of words recalled in a task in which a short list of words list presented for recall in correct serial order depended on the spoken duration of the words in the list, suggesting that time limits are an important component of STS. On the other hand, Glanzer and Razel (1974) found that subjects recalled approximately the same number of items in a STM task regardless of the length of the items (words vs idioms), supporting STS capacity limits. One reason for these contradictory findings may be the use of different tasks. Baddeley, Thomson, et al. (1975) used serial recall of short lists of words that were from a small pool of items sampled with replacement. Glanzer and Razel used the immediate free-recall task in which subjects recalled longer lists of words selected without replacement, so words were never repeated across lists. Time limits may be most important when rote rehearsal processes are used to maintain information: However, when such strategies are less useful, a limitation based on the number of items being represented may be more important.

Transfer of Information to Long-Term Store

In multi-store models, one of the primary functions of STS is to hold information until it can be represented in a more permanent state. Because rehearsal was assumed to be the critical mechanism for this transfer, many researchers attempted to distinguish between STS and LTS by examining the relationship between rehearsal and long-term retention. One line of evidence suggesting that the formation of permanent memory traces depends on rehearsal of information in STS came from the U-shaped immediate free-recall curve (Glanzer & Cunitz, 1966). Typically, recall of the first few items from a list is very high compared to items that occur in the middle of the list. This finding is known as the primacy effect. Glanzer and Cunitz believed that primacy reflects the effect of rehearsal, and studies show that initial items do receive more rehearsal than items from other positions in the list (Rundus & Atkinson, 1970). Recall of the last few items from a list is also very high, even though those items do not receive many rehearsals.

Glanzer and Cunitz argued that this recency effect occurs because the last items have not yet decayed from STS and, therefore, are easily retrieved. In an attempt to show that the recency effect reflects short-term storage, they manipulated the delay between the presentation of the list and the recall of list items. During the delay, subjects counted aloud from a target number until they received a cue indicating that they should recall the items. As expected, this delay reduced the recency portion of the curve and left primacy unaffected. As we will see, however, performance on the recency portion of the immediate free-recall curve[4] does not seem to have any relationship to other measures of short-term memory, so it may not be a *valid* reflection of short-term memory.

Coding

Another distinction between STS and LTS, made by structural theorists, involved the type of coding used in each store. Early evidence suggested that information in STS is coded phonologically,[5] whereas information in LTS is coded semantically. Again, such a dissociation between the two systems supports the existence of two distinct memory structures. For example, in a memory span task, subjects recalled fewer letters when lists were made up of phonologically similar letters (e.g., *C, B, G, V, T, P*) as compared with lists of phonologically dissimilar letters (e.g., *F, J, M, R, L, Y;* Conrad & Hull, 1964). Errors made on similar lists are typically phonological confusions (e.g., responding with *B* rather that *D*), suggesting that items are coded by their sound even if they are presented visually. Although the phonological similarity effect supports the use of phonological codes in STS, it does not address whether or not semantic, or any other types of coding, can also be used in STS. Baddeley (1966) presented similar and dissimilar lists to subjects for serial recall, but in his study some lists included either phonologically similar or dissimilar words, whereas others included either semantically similar or dissimilar words. Phonological similarity decreased recall performance, but semantic similarity only slightly affected recall performance.

Kintsch and Buschke (1969) used a serial probe task similar to the one used by Waugh and Norman (1965). They found that phonological similarity affected recall for targets at the end of the list (putatively items in STS), but not at the beginning of the list (items in LTS). Further, semantic similarity reduced recall of targets at the beginning of the list, but not at the end of the list. If the recency portion of the list reflects information in STS and the primacy portion reflects information that has been stored in LTS, as already described (Glanzer and Cunitz, 1966), then finding phonemic similarity effects at the end of the list and semantic similarity effects at the beginning of the list supports differential coding and the dual-store approach. That does not mean that all temporary traces, however, are necessarily coded in the same format. It is likely that different tasks used to study human memory encourage or force the use of different codes to represent the information.

Neuropsychological Evidence

In addition to the behavioral data, neuropsychological case studies provided converging evidence for multi-store models. The well-known case of HM suggested that different brain areas may underlie the functioning of proposed memory stores, and thus it supported a multi-store system (Milner, 1966).[6] HM showed normal performance on short-term memory tasks and good recall of events that occurred prior to his surgery. The deficit was clearly in the storage of information in the long-term store. Shallice and Warrington (1970) presented the case of a subject who appeared to be the complement of HM. Their patient showed impaired performance on short-term memory tasks and normal performance on long-term memory tasks. Baddeley and Warrington (1970) compared patients with Korsakoff's syndrome (amnesia resulting from chronic alcoholism) to normal subjects on a variety of short-term and long-term memory tasks. The amnesics showed normal forgetting on the Brown-Peterson task and normal recency in free recall but reduced primacy as compared to control subjects.[7] Together, these results suggested that two separate memory structures, in fact, exist and that performance on short and long-term memory tasks are mediated by different brain areas.

There is now strong evidence that the hippocampus is involved in the temporary storage of information and that bilateral removal of the hippocampus, as in HM, leads to reduced ability to recall information presented more than a few minutes earlier (Kolb & Whishaw, 1990). This reduced ability to store information so that it can be easily retrieved later appears to be specific to what some theorists refer to as *explicit memory*.

HM, for example, learned the backward mirror-tracing task and retained the ability from day to day. Thus, although the research on the hippocampus supports a multi-store approach, it also suggests that the ideas hold for only certain kinds of representations or types of retention test.

Problems with the Multi-Store Model

Support for multiple memory structures based on both behavioral and neuropsychological dissociations seemed fairly strong; STS and LTS appeared to have distinctly different characteristics. However, a great deal of criticism surrounds research that distinguished the characteristics of short-and long-term memory as well as some of the basic features of the multi-store model (cf., Crowder, 1983, 1993). Crowder argued that in postulating distinct memory stores, multiple-memory approaches typically confounds codes and processes.[8] More specifically, he suggested that the distinction between STS and LTS may simply be a distinction between phonetic coding and semantic coding. These codes may be used in different tasks and may affect characteristics of retention, but it is not necessary to hypothesize that different processing systems are involved. In support of this idea, Crowder suggested that characteristics that appear to be unique to STS can be accounted for within a unitary memory framework. For example, as previously discussed, forgetting in the Brown-Peterson task may depend on interference and not on a decay mechanism (Keppel & Underwood, 1962). Similarly, the recency effect found for immediate free recall is also found in long-term memory tasks (Bjork & Whitten, 1974), as is the phonemic similarity effect (Gregg & Gardiner, 1984).

An elaboration on Crowder's point is that much of the early work on STS and, indeed, more recent work on the articulatory loop, made use of a procedure in which subjects were presented a short list of words or letters for serial recall. Further, the list items were generally selected from a small pool of items with replacement from list to list. It is likely that procedures such as the immediate free recall of 12- to 15-item lists in which the items are never repeated gives rise to very different coding and processing. Thus, it is difficult to know how generalizable results are from one task to another, even though both are putative short-term memory tasks.[9] We will further discuss *measurement* aspects of short-term memory in a later section.

Another problem with the multi-store model is the assumption of serial processing through the different stages. First, serial ordering from STS to LTS does not make logical sense. Information is supposed to be held in STS before making contact with LTS, but how can that information be identified without relying on LTS? If you present me with the word *dog* as one item in a list to recall, I certainly recognize the word and have associations with its meaning soon after I see or hear it. Therefore, it must have led to access of some trace in the LTS prior to maintenance in STS. In addition, Shallice and Warrington (1970) pointed out that case studies of patients who have STS deficits without LTS impairments are inconsistent with the serial processing through STS and then LTS. If these processing stages are serial, then deficits in STS should also lead to deficits in LTS because long-term storage is supposed to be dependent on processing information through STS. To circumvent this problem, Shallice and Warrington suggested that information makes contact with STS and LTS in parallel. Most modern theories of short- and long-term memory have eliminated the assumption of serial processing through different stages and assume that STS is a more highly activated subset of knowledge units in LTS.

Further, it is important to think about what really is being learned in a short-term memory task. What is retained in STS is not my knowledge about *dog* but that *dog* was on the list I just saw or heard, that it was near the middle of the list, that it followed the word *table,* and that it, in turn, was followed by the word *aardvark*. That would be the knowledge transferred to LTS or learned through appropriate coding or rehearsal. Correspondingly, that is the knowledge that would be lost either through decay, or interference according to more recent multi-store theories.

Finally, one of the biggest concerns about the multi-store model involved the assertion that strength of the representation in long-term memory is a function of the amount of rehearsal in STS. Rundus and Atkinson (1970) did show that immediate recall is a function of the number of rehearsals given to an item. However, other studies showed that delayed recall of information is not a function of the amount of simple rote rehearsal. For example, Tulving (1966) found that recall of a list of words was not enhanced when subjects repeated the list aloud six times. Similarly, the results of a study by Craik and Watkins (1973) argued against the idea that time spent in STS determines the transfer of information to LTS. Subjects received a target letter and then heard a list of words. They were to retain in memory the first word that began with the target letter until another word beginning with that letter occurred. At this point, subjects could drop the first word and hold the second word in memory. The task continued in this manner until the end of the list, when subjects reported the last word in the list that began with the target letter. By manipulating the number of words between target words, Craik and Watkins effectively manipulated the amount of time a given word spent in STS. At the end of the experiment, subjects were asked to recall all of the critical words from the experiment. Craik and Watkins found no relationship between time spent in STS and the probability of long-term recall. Because the time spent in STS should reflect the amount of rehearsal an item receives, these studies suggested that rehearsal alone is not sufficient to account for the transfer of information to LTS. Problems with serial processing, rehearsal as a mechanism of information transfer, and questions about the distinctiveness of the two stores all contributed to the decline of the popularity of structural models of memory.

In response to many of the problems with multi-store approaches, Craik and Lockhart (1972) suggested that apparent differences in STS and LTS reflect differences in the processes used to analyze information. For example, differences may result because of differential processing demands for tasks used to examine STS and LTS. According to this approach, the time spent in STS is not the factor that determines the strength of the memory trace. Rather, the *level of processing* may be the critical determinant of long-term storage, with "deeper" processing yielding stronger traces. Inconsistent results found in studies examining the relationship between rehearsal and transfer to LTS, then, may be the result of different types of rehearsal processes.

Although Craik and Lockhart (1972) did not propose the existence of multiple memory systems, they did distinguish between information that is currently in consciousness and that which is not. They assumed that a limited capacity central processor (i.e., controlled attention), which operates to provide the different levels of processing, is responsible for keeping information in consciousness and that information which is in consciousness is easily and accurately retrieved. When attention is directed away from information currently in consciousness, that information decays and is replaced by new information. Craik and Jacoby (1975) pointed out that the Craik and Lockhart view is not unlike the STS proposed by many multiple memory theorists (e.g., Atkinson & Shiffrin, 1968; Norman, 1969). The difference between the two approaches is that Craik and Lockhart did not believe that the properties of stored information depend on the structure in which information is held; nor did they believe that information is transferred between different stores. Consciousness, or *primary memory,* as Craik and Lockhart prefer to call it, is not a structure; rather, it is the activation of processes used to analyze information. Limitations in the amount of information that can be "in" primary memory are due to attentional processing limitations, not structural limitations. As you will see, this view is similar to the concept of the central executive proposed by Baddeley and Hitch (1974) and the idea of working memory capacity proposed by Turner and Engle (1989) and Just and Carpenter (1992).

Working Memory

A number of intellectual influences served to move thought about short-term memory to what is now called *working memory.* One of these influences was the resurgence of Piagetian constructs relating temporary storage and controlled attention in the developmental psychology literature. These are most clearly seen in the work by Pascual-Leone (1970) and Case (1985), who proposed the notion of *M-space,* which is similar to what later would be called working memory.

In the experimental psychology literature, Baddeley and Hitch (1974) proposed a flexible and more complex temporary storage system that avoids some of the weaknesses of the multi-store model. A major liability for the multi-store model is evidence that neuropsychological patients with STS deficits do not have impaired performance on complex cognitive tasks (e.g., Shallice and Warrington, 1970; Warrington, Logue, & Pratt, 1971; Warrington & Weiskrantz, 1972). This finding posed a problem for the multi-store model because STS is viewed as the limiting factor in the information-processing system, and so STS impairments should lead to impairments in other types of cognitive processing. In a similar vein, if STS is the bottleneck in processing, then processing on a task that depends on STS should be less efficient when STS is occupied with another task. Baddeley and Hitch examined this hypothesis in a series of studies that manipulated memory load during complex task performance. They showed that although performance on reasoning and comprehension tasks was not affected by a small memory load of two to three digits, performance on these tasks did decrease with a load of six items. These results led Baddeley and Hitch to propose a system of working memory (WM) consisting of three components (see also, Baddeley, 1986, 1996; Baddeley & Hitch, 1994).[10] Two of these components, the *articulatory loop* and the *visual-spatial sketchpad,* are "slave" systems, and the other is a *central executive.* The slave systems are largely responsible for the maintenance of acoustic and visual information, whereas the central executive is responsible for control of information processing. Baddeley and Hitch argued that WM is a unitary system limited in both storage and processing and that there is some flexibility in the allocation of attention to the components. This flexibility accounts for the results of their studies, which found that articulatory suppression and concurrent memory load interfere with the performance of learning and reasoning tasks, more or less independently, depending on the extent to which the task requires the use of phonological information and the extent to which it requires controlled attention for processing. Similarly, their model can account for neuropsychological case studies of patients who have STS deficits but who do not have serious deficits on many complex cognitive skills, by assuming that these individuals have damage to one of the slave systems rather than to the central executive.

Slave Systems

Most of Baddeley's work has focused on the two slave systems, with the articulatory loop receiving more attention than the sketchpad. As with the early work on STS, examinations of WM first focused on issues of coding, rehearsal, and loss of information. Thus, like early models of STS, the two slave systems represent the more rigid, structural aspects of Baddeley and Hitch's (1974) working memory model.

The articulatory loop most closely resembles earlier conceptions of STS because it consists of a limited duration, speech-based representation and is dependent on articulatory rehearsal for the maintenance of information. As previously described, reduced recall with phonemically similar lists suggests that information held in STS may be coded phonologically (Conrad, 1964; Conrad & Hull, 1964). Baddeley and Hitch suggested that because the articulatory loop is responsible for the temporary storage of verbal information, it is the locus of phonemic similarity effects. Further evidence for speech coding and time limits on the articulatory loop comes from studies showing that word length affects recall. That is, recall in a short-term memory task depends on the time it takes to articulate those words; fewer items from lists of long words are recalled than from lists of short words (e.g., Baddeley, Thomson, et al., 1975; Case, Kurland, & Goldberg, 1982; Ellis & Hennelley, 1980; see also, Cowan et al. 1992; Cowan, Keller, et al. 1994; Cowan, Wood, & Bourne, 1994). Further, the importance of articulatory rehearsal for the maintenance of verbal information is demonstrated by the fact that the phonemic similarity effect and word length effect are eliminated under conditions that prevent rehearsal. For example, both effects are eliminated by articulatory suppression, a procedure in which rehearsal is prevented by having subjects vocalize a sequence (e.g., saying "blah, blah, blah") while completing an immediate recall task. The elimination of the effects by articulatory suppression suggests that these effects are the result of articulatory rehearsal (Baddeley, Lewis, & Vallar, 1984; Baddeley, Thomson, et al. 1975; Levy, 1971).

If the phonological loop is to be considered a viable component within the working memory system, it must have functional significance beyond its role in simple short-term memory tasks. Because the phonological loop stores verbal information, it is not surprising that researchers have considered its role in both language comprehension and language acquisition. Results concerning the relationship between the phonological loop and language comprehension have been mixed (Baddeley & Wilson, 1985; Butterworth, Campbell, & Howard, 1986; Caplan & Waters, 1992; Wilson & Baddeley, 1993). However, developmental and neuropsychological studies suggest that the phonological loop may be an important mechanism in language acquisition (Baddeley, Papagno, & Vallar, 1988; Gathercole & Baddeley, 1989, 1990; Papagno, Valentine, & Baddeley, 1991; Papagno & Vallar, 1992).

The other slave system proposed by Baddeley and Hitch (1974), the visuospatial sketchpad, has received far less attention than the phonological loop. The sketchpad functions as a temporary store for holding and manipulating visual and spatial information. Current controversy over the visuospatial sketchpad centers on whether or not there are two interactive visual systems, one that holds visual patterns and the other of which represents spatial information. Both behavioral and neuropsychological dissociations support the dissociation of the sketchpad into two interactive subsystems (Baddeley Grant, Wright, Thomson, 1975; Baddeley & Lieberman, 1980; Farah, 1984; Farah, Hammond, Levine, & Calvanio, 1988; Hanley, Young, & Pearson, 1991; Logie, 1986).

Central Executive

The central executive component of Baddeley and Hitch's (1974) model is the least specified of the subsystems. It was conceived of as a limited capacity processor that is flexibly allocated to processing and/or storage functions. Baddeley (1986, 1996) suggested that the central executive is similar to the concept of *supervisory attentional system* (SAS) proposed by Norman and Shallice (1986). According to Norman and Shallice, actions are carried out via the activation of schemas. Given the appropriate goals and stimulus context, a schema will be automatically activated and initiate a sequence of actions. It is possible that several schemas may be carried out simultaneously (e.g., walking and talking at the same time), but sometimes schemas will conflict with one another or will need monitoring to catch errors. Norman and Shallice proposed two levels of control to activate schemas and resolve conflicts. First, the current goals can lead to the enhanced activation of some schemas and inhibited activation of others, a step that serves to select the most appropriate, or highly activated, schema for action. This complementary process is a result of automatic spreading activation and thus does not require attention unless the procedure is error prone and must be monitored. Second, a limited-capacity attentional control system, the SAS, mediates the scheduling of contending schemas. The SAS is necessary when activated schemas are incompatible with current goals.

Baddeley (1986) suggested that the SAS could account for the results of Baddeley and Hitch's (1974) studies, which showed that some primary tasks were affected by concurrent load while others were not. In several studies Baddeley and Hitch found that concurrent digit load did not affect retrieval accuracy, but did affect retrieval latency. Baddeley argued that neither retrieval nor maintenance of the digit load involves heavy demands on the SAS, so both tasks can be performed accurately. In contrast, retrieval latency was affected by load because performance on the two tasks simultaneously requires more time-consuming contention scheduling than when one task is performed alone. Another set of studies showed that generation of exemplars from a category (e.g., animals) was affected by load. Baddeley reasoned that there are no schemas that can be automatically activated for category generation, so this task relies heavy on the SAS.

Baddeley's reliance on the Norman and Shallice (1986) model to further specify the central executive makes it apparent that the central executive is really a method for allocating attention. So, this aspect of Baddeley's model is similar to Broadbent's (1958) information-processing model in the sense that attentional and short-term storage systems are separated, but highly interconnected.

Baddeley has recently attempted to identify processes that are characteristic of the central executive and has examined executive control by studying neurological patients who appear to have deficits in cognitive control. A series of studies showed that Alzheimer patients exhibited impairments in coordination of slave systems, switching retrieval plans, selective attention, and activation of long-term memory (e.g., Baddeley, Logie, Bressi, Della Sala, & Spinnler, 1986; Baddeley & Wilson, 1988; Baddeley, Bressi, Della Sala, Logie, & Spinnler, 1991). As will be shown, there is now considerable evidence for a view that areas of the frontal lobes, particularly the prefrontal cortex, are important to central executive functioning.

Baddeley and Hitch (1974) set out to create a model of short-term storage that was more flexible than traditional short-term memory systems. According to their model, WM has both storage and processing functions. The slave systems are responsible for the maintenance of information and are similar to traditional STS models. Perhaps because of a rich history of research on temporary storage, these components are fairly well specified and allow for the generation of testable predictions. The central executive is an attentional control system and provides a great deal of flexibility to the WM system. Although the central executive is far less specified than the slave systems, tying it to short-term storage represents an important step toward the understanding of interactions between attention, short-term storage, and complex cognition. Because the model incorporates both processing and storage functions, it can be considered more dynamic than structural models proposed in the past, but it is still structural in nature and it is still a memory system that is distinct from other processing systems.

We should also point out that the two so-called slave systems of the Baddeley and Hitch model are not necessarily structural in nature. It is probable that the human brain has evolved to be especially good at processing speech and visual-spatial information. However, it is also possible to think of articulatory and visual-spatial coding as simply two of myriad other possible ways of coding information. For example, Reisberg, Rappaport, and O'Shaugnessy (1984) taught subjects to code lists of digits by tapping the corresponding fingers and found that the articulatory loop and the "finger" loop are independent of each other. These findings support the notion that the human information-processing system is very flexible in how it can represent information and that each format may have distinctively different characteristics.

Other Approaches to Working Memory

In contrast to Baddeley's (1986) structural approach to working memory, other researchers have described more dynamic or process-oriented models of working memory. Anderson (1983a) proposed a model of cognition in which working memory is simply the set of long-term memory units currently activated above a critical threshold. Schneider and Detweiller (1987) proposed a parallel distributed model of working memory, which, like Anderson's (1983a) model, views short-term storage as activated information. They argued that buffer models have difficulty explaining why temporary interruptions do not destroy performance on a complex task and force the individual to start the task over. In fact, we humans have rapid access to a great deal of information. For example, as I sit here writing this passage, I can quickly think about the previous section I wrote or what I want to say next. If I get interrupted by a student knocking at my door, I can go to the door, open it, greet the person, answer a question, talk about how the student is doing in my class, and so on, and quickly return to the point in my writing where I was interrupted. Miller's (1956) notion of 7 ± 2 items in temporary storage does not say that we can retain 7 ± 2 items in a number of different domains or contexts, but our common experience is that, in fact, we can. The Schneider and Detweiller model is unique in attempting to address this quality of working memory.

In their model, the information-processing system consists of three levels of analysis within many different processing modules. A module is analogous to a brain region that specializes in a class of processing (e.g., visual, motor, or lexical processing). Output units from each module may become activated and transmit activation to other modules or to the next level of analysis. Information in working memory is activated above a critical threshold. There is an active buffer, or short-term store, for each modality or stimulus context. Thus, my writing buffer would not necessarily be interfered with by my student-asking-question buffer. Schneider and Detweiller do not make qualitative distinctions between short and long storage of information; rather, in their model, short- and long-term memory are differentiated by temporal distinctions. So, short-term storage is limited in duration, whereas long-term storage is not.

In addition, there are limitations in the amount of activation that can be transmitted at a given time; these limitations are analogous to capacity limitations in buffer models. Some information can be activated and transmitted automatically, without consuming limited controlled-processing resources. But, control processes at each level of analysis determine what information will and will not be transmitted and the order of transmission. The functions of these control processes are akin to the functions of Baddeley's central executive system; however, Schneider and Detweiller do not invoke a singular executive component that controls processing. Rather, control is distributed among modules and levels of processing within modules. Further, they suggest that there is not a central capacity limitation; processing limitations will depend on the specific modules being used. The most apparent benefit of Schneider and Detweiller's (1987) model is that it allows for a great deal of interaction among different types of processing; thus, it has more power to explain complex processing than do simple buffer models. However, along with the complexity comes a lack of specificity; it is difficult to generate testable hypotheses from their model. Further, there is growing neurological evidence of a central attentional system that is not domain-specific (Posner & Peterson, 1990).

Another comprehensive model of the attention and memory system was proposed by Cowan (1988, 1995). He proposed a dynamic model of attention and memory in which WM is considered an activated subset of LTM. In fact, he suggested that Anderson's (1983a) and Schneider and Detweiller's (1987) models are compatible with his ideas, but he believed that they reflect a different level of analysis. Cowan proposed that there is one storage system that consists of elements and their features in long-term memory (e.g., acoustic, visual, semantic, etc.). At any given time these elements may exist in one of three states of activation. First, they may be inactive, a state that is akin to long-term storage. Next, they may be in a moderately activated state, but outside the "focus of attention." Information in this state is outside conscious awareness, but can influence processing (e.g., semantic priming, subliminal perception, implicit memory). This level of activation represents the passive storage of information in Baddeley's slave systems. The duration of this activation is limited, and without strategies for maintaining activation, information will quickly return to an inactive state. Finally, when information is attended, the level of activation will increase and the information will become the focus of attention, which represents the highest level of activation and is synonymous with conscious awareness. Cowan suggested that there are severe limitations in the amount of information that can be maintained in the focus of attention and that these limitations reflect the capacity of the central executive in Baddeley's model. Cowan eliminated the need for multiple memory structures by invoking different levels of activation that have distinct limitations. Further, he suggested that although much of the evidence supporting short- and long-term stores may actually reflect different processes used for activating information and maintaining activation, there is enough other evidence that supports a qualitative distinction between short- and long-term storage systems to warrant a multiple memory system approach. Cowan suggested that viewing short-term storage as the activation of long-term memory units represents a middle ground between approaches that posit multiple memory structures and completely unitary approaches. Activation approaches eliminate the need for completely separate memory systems while continuing to suggest that short- and long-term storage obey different sets of rules.

In addition to the memory system, Cowan's model includes a limited capacity central processor, or central executive, that is responsible for all control processes. Cowan defined control processes as those processes that are under voluntary control and that require attention for implementation (see Kahneman, 1973; Posner & Snyder, 1975; Shiffrin & Schneider, 1977). Cowan also solved the question of whether forgetting from short-term memory is based on time and decay or capacity limits and displacement. He assumed that the focus of attention is limited in capacity and that newly focused information displaces old information. Information that is activated but outside the focus of attention is lost over time through decay.

The Anderson (1983a), Schneider and Detweiller (1987), and Cowan (1995) views are similar in proposing an active portion of memory. They are different in how they handle the notion of executive control. Anderson proposes controlled attention for maintaining the activation of goals and for resolving conflict, which is similar to Shallice's SAS. Schneider and Detweiller propose that attention, like the memory representations, is distributed across modules. Cowan's view of executive control is, like Anderson's, more in line with Shallice's and with the Baddeley and Hitch central executive.

Working Memory Capacity and the Central Executive

Just and Carpenter (1992) and Engle, Cantor, and Carullo (1992) proposed models to explain individual differences in working memory. As with Anderson (1983) and Schneider and Detweiller (1987), these models assumed that working memory is an activated subset of long-term memory and that individuals differ in the amount of activation available for processing. Work on individual differences in working memory stems from studies that show that unlike traditional short-term memory measures, measures of WM correlate with performance on complex cognitive tasks. For example, Daneman and Carpenter (1980) showed that a measure of WM capacity that involves both storage and processing is related to performance on a higher-level cognitive task. In their reading span task, subjects read a series of sentences and tried to remember the last word of each sentence. The maximum number of final words recalled correlated with reading comprehension and verbal scholastic aptitude scores (VSAT). Other researchers have shown that the reading span and similar WM measures are highly correlated with complex cognitive measures such as writing ability, following directions, logic learning, and vocabulary learning (e.g., Benton et al. 1984; Daneman & Carpenter, 1983; Daneman & Green, 1986; Engle, Carullo, & Collins, 1991; Kiewra & Benton, 1988; Kyllonen & Stephens, 1990). An understanding of why these measures are related and what components of the WM span task are important to individual differences in complex task performance will help to specify those aspects of working memory that are generalizable to real-world cognition and thus are less likely to be peculiar to a specific experimental task.

Daneman and Carpenter (1980, 1983) and Just and Carpenter (1992) argued that individual differences in the level of activation are important to lower-level language processing, such as syntactic parsing and disambiguation of ambiguous linguistic units.[11] They argued that these differences reflect differences in activation available to language processing, but, like Schneider and Detweiller, they assumed domain-specific limitations on activation.

On the other hand, Turner and Engle (1989) and Engle et al. (1992) assumed the central executive reflects a domain-free attention capacity limitation and suggested that individuals with high and low working-memory spans should differ on capacity-demanding tasks, no matter what the processing domain. Like Daneman and Carpenter (1980), they showed that reading-span task performance correlated with global verbal comprehension measures (e.g., VSAT). However, Turner and Engle (1989) also showed that performance on a working memory task that required subjects to solve math problems while remembering words correlated with verbal comprehension as well as did the reading span measure. This finding suggests that the storage component of these span tasks reflects differences in a central attention limitation, not differences due to task-specific processing.

Engle et al. (1992) suggested that WM span differences reflect differences in the overall amount of activation available to the WM system for processing. Individuals with greater WM capacity are able to maintain activation of more LTM knowledge units than individuals with less WM capacity. Thus, high–WM span individuals can keep more units in an active state and available for rapid retrieval and further processing. In addition, Engle, Cantor, and Carullo argued that the amount of activation available is a stable characteristic of the information-processing system and that it changes little with changes in knowledge structure. Cantor and Engle (1993) provided evidence for the general capacity model by showing that working memory capacity and measures presumed to reflect activation of information in LTM are related. In that study, high and low working-memory span individuals learned a series of unrelated sentences. Each sentence consisted of a subject and a predicate (e.g., "The lawyer is in the park"). Further, each subject was paired with more than one predicate (e.g., *lawyer* might be paired with *park* and *boat*). Next, the participants performed a speeded recognition task in which they determined whether sentences belonged to the studied set or not. Typically, in this task, reaction time increases as a function of fan size (fan size is the number of times a given subject appears in the stimulus set with a different predicate; Anderson 1983b) and the increase is attributed to the division of activation among a greater number of knowledge units for sentences with larger fans. Cantor and Engle found that reaction time increased across fan size for both high-and low-WM groups. However, the increase was more dramatic for the low-span subjects, suggesting that it took them longer to activate target sentences because they had less overall activation spreading in the network than did high-span subjects. In fact, when the slope of the fan effect was partialled out of the correlation between span and verbal abilities, this correlation was no longer significant, suggesting that long-term memory activation is an important component of the relationship.

The general capacity model was elaborated by Conway and Engle (1994). They studied whether individual differences in working memory reflect differences in automatic spreading activation, as was argued in Engle, Cantor, and Carullo (1992), or differences in controlled attention, that is, the central executive. Conway and Engle suggested that Cantor and Engle's (1993) task is not sensitive to a distinction between controlled attentional processing and automatic activation, because the fact retrieval task used by Cantor and Engle requires both processes. In the verification phase of the 1993 task, subjects might encode the subject of the target sentence and activate all associated predicates. Next, subjects searched the activated set of information and determined whether the target was a member of the set. Because each sentence in the Cantor and Engle study shared a predicate with another sentence (e.g., "The lawyer is in the boat" and "The teacher is in the boat"), it is possible that there was response competition or conflict that would necessitate a controlled search of the activated information. Thus, differences between high and low WM groups could reflect differences in automatic activation of LTM units, a controlled search of these activated concepts, or differences in both processes. In order to assess differences in these two types of processing, Conway and Engle had subjects retrieve facts from either active or inactive memory.

High and low working-memory span subjects learned an association between items in memory sets of various sizes and a digit set cue that corresponded to the size of the set. Each set contained unique items in one experiment and overlapping items in another. Next, subjects performed a speeded verification task in which they saw a digit and an item and pressed a key to indicate whether the item was a member of the set. In the short-term memory condition, the digit set cue appeared 1 second before the probe. Thus, subjects knew which memory set was being tested and could retrieve the set information into active working memory before the probe appeared. Therefore, recognition only depended on a search of short-term memory. In the long-term memory condition the set cue and probe appeared simultaneously, so the set had to be activated before it could be searched. It was assumed that, in this condition, the subject would need to access the set information from long-term memory, move it into short-term memory, and then do a search of short-term memory. Thus, the two conditions differed in that a retrieval from long-term memory was necessary for the latter condition. Conway and Engle found that the slope of the set size function did not differ for high- and low-span subjects in the no-interference condition, that is, the condition in which there was no overlap in set membership. In the interference condition with overlapping sets and presumed response competition, the slope o the set size function *did* differ for high- and low-span subjects. Unde interference conditions, the low-span subjects were much slower to iden tify the item as belonging to the set. Conway and Engle (1994) argue that this difference between the search functions of high- and low-spa subjects resulted from search of activated memory, because the no dela and 1-second delay conditions showed the same slope.

For high-WM subjects, verification times across memory set size wer the same for overlapping and unique memory sets. In contrast, set overla slowed the verification times for low-span subjects. In summary, whe unique sets were tested, and thus no response competition was present the slope functions for high- and low-WM subjects were nearly identica However, low-span subjects' performance was slowed by overlap be tween memory set items, but the performance for high-span subjects wa *not*. Conway and Engle argued that the verification task in the condition with nonoverlapping sets could be performed on the basis of automati spreading activation between the digit probe and the target. In the over lapping sets conditions, however, a conflict would arise when the activa tion from the target would spread to the correct probe *and* to anothe probe, hence, a condition that would require the supervisory attentiona system of Norman and Shallice. Conway and Engle further argued tha this conflict forced the low-span subjects to do a controlled, serial searc in that condition. The high spans were argued to use their greater atten tional resources to suppress the irrelevant link so they did not need t do a controlled search of the list.

Conway and Engle (1994) argued that working memory capacity i important when retrieval is necessarily or voluntarily achieved throug a controlled search, but not when achieved through passive automati activation. This conclusion was also supported by a set of studies by Ro sen and Engle (1997) in which high and low working-memory subject were instructed to generate as many different animal names as possibl over a 10-minute period. In the first study, high-span subjects generate about 40% more animals than low-span subjects. One possible explana tion was that low-span subjects relied largely on automatic spreading act vation for retrieval but that high-span subjects used controlled attentio for search and to suppress previously retrieved responses. This explana tion was supported by the finding that doing the retrieval under the workload of an attention-demanding detection task hurt performance for the high-span subjects but had no effect on the retrieval of the low-span subjects.

On the basis of these studies we have proposed that individual differences on measures of working memory capacity reflect differences in controlled attention capability and that those differences will only be reflected in situations that either encourage or demand controlled attention (Conway & Engle, 1994; Engle, Conway, Tuholski, & Shisler, 1995; Rosen & Engle, 1997a). Controlled attention is necessary when task goals may be lost unless they are actively maintained in working memory, where actions contending for the same stage must be scheduled, where conflict among actions must be resolved, where there is value in maintaining some task information in the face of distraction and interference, and where there is value in suppressing task-irrelevant information.

Measurement Issues with Working Memory

As we have traced the history of the literature on short-term and working memory, we have seen that the models have become more complex and more flexible. We have also seen a divergence of methodology from strictly experimental studies to studies of individual differences on hypothesized constructs. This evolution has given rise to some concerns that are novel to most experimental psychologists. Experimental psychologists have, traditionally, not paid much attention to measurement aspects of the tasks they use. Psychometric issues such as task reliability and validity are more often considered in applied areas or areas such as social and personality psychology. Many tasks have been used to study short-term memory and working memory, and because individual differences in working memory have become important, concerns about psychometric issues have increased. Validity as an issue is reflected in such questions as: Does the task measure what you want it to measure? Do different measures of working memory reflect the same construct? Do tasks that putatively measure short-term memory measure the same construct as tasks that putatively measure working memory? and Do working memory and/or short-term memory have construct validity? or Does the construct have some importance or relationship to real-world behavior?

In a review of research on memory span, Dempster (1981) argued that if short-term memory is important to real-world tasks such as reading, and if memory span is an index of short-term memory, memory span should correlate with measures of reading. However, simple span measures, the digit span in particular, do not consistently correlate with measures of reading comprehension. Dempster argued that part of the problem is that simple memory-span scores are simply not very reliable: the same subject can show wide differences in span when different measures are used. The simple word span is typically more reliable than the digit span, which might explain the fact that nearly all of the studies from our lab (Engle et al., Collins, 1991; LaPointe & Engle, 1990; Turner & Engle, 1989) show that simple word span correlates with reading comprehension. We have also found that the digit span does not correlate well with other measures of short-term memory; again, the failure to find such correlations consistently is probably a reliability problem. So, from a simple measurement standpoint, the digit span is probably not a very sensitive measure and, thus, is not a very useful measure of short-term memory; the word span may be a better measure.

Another important measurement issue is whether STM and WM tasks reflect the same underlying construct. We recently addressed this question in a large-scale factor analysis (Engle, Tuholsky, Laughlin, Conway, 1998). The study included tasks traditionally thought of as short-term memory tasks, including forward word span with phonologically dissimilar words, forward word span with phonologically similar words, and the recency score from immediate free recall. Other tasks were chosen to reflect working memory capacity or the central executive. These included the random generation task (Baddeley, 1996), reading span (Daneman & Carpenter, 1980), operation span (Turner & Engle, 1989), and two tasks from the CAM4 battery: continuous opposites and ABCD (Kyllonen & Christal, 1990). Another set of tasks was used because whereas some authors have referred to them as short-term memory tasks, others have referred to them as working memory tasks. These include the backward word span with phonologically dissimilar words, keeping track task, counting span, and recall from all the serial positions in immediate free recall except recency. In addition to the memory tasks, subjects were tested on the Raven's Progressive Matrices Test and the Cattell Culture

Fair Test, both of which are nonverbal tests of general fluid intelligence (*gF*). We tested 133 subjects individually over three sessions and obtained the Verbal and Quantitative Scholastic Aptitude Test scores for them all.

We performed a series of exploratory then confirmatory factor analyses on the memory tasks. These analyses showed that two different factors were necessary to account for the variance in the memory task scores. What we called the short-term memory (STM) factor included the two forward span tests and the backward span. The working memory (WM) factor included operation span, reading span, counting span, keeping track, secondary memory component from immediate free recall, and the two tasks from the CAM4. A partial regression analysis showed that when STM was controlled for, WM correlated with the general fluid intelligence tests, $r = .5$. However, if WM was controlled for, STM did not significantly correlate with *gF*. This finding shows quite conclusively that the STM tasks and the WM tasks reflect different underlying constructs. Both factors, however, contributed significant and independent variance to the Verbal SAT, a measure of verbal skills, including reading comprehension. Our present thinking about these results is that whereas the WM factor reflects controlled attention or attentional resources capability, the STM factor may reflect some basic aspect of speech representation that is also important to the VSAT. An appealing possibility is that the WM factor represents the central executive component of the Baddeley and Hitch (1974) model and the STM factor represents the capability of one of the slave systems, namely, the phonological loop. If this interpretation is correct, then tasks could also be used that reflected the domain-specific capability of the visuospatial sketchpad as well as the phonological loop.

The random generation and recency portion of immediate free recall did not fit with either factor and were dropped from the analyses. This finding calls into question the use of the recency portion of the free recall task for making inferences about short-term memory.

The results of the study fit nicely with work by Kyllonen and Christal (1990), which argues that working memory capacity is an important component of what is commonly thought of as general fluid intelligence. As we will see in the next section, the Kyllonen and Christal conclusions also tie in with work on controlled attention and the functions of the frontal lobes of the brain.

At this point, it might be useful for us to attempt some generalizations about short-term memory and working memory. We would argue, following Craik and Jacoby (1975) and Cowan (1995), that many different processes can be used to maintain the temporary activation of memory units. Certainly phonological, visual, and spatial coding reflect common means of coding, but there are undoubtedly other means as well. It seems reasonable to think of those knowledge elements activated above some threshold as reflecting the contents of short-term memory. Some small number of those activated units may be in the focus of attention. From a measurement standpoint, individuals almost certainly differ in their ability to use speech-based, spatial, and visual coding, and those would be reflected in measures appropriate for each code. The vast majority of short-term memory studies reflect the use of phonological coding of short lists of words and rote rehearsal, using that code, to maintain activation of the representation. As Cowan (1995) pointed out, not all information in short-term memory, that is, activated knowledge, can be the focus of attention, because the focus of attention is quite limited in capacity. We argue that focused or controlled attention corresponds to Baddeley and Hitch's central executive. Individual differences in "working memory capacity" (Just & Carpenter, 1992; Turner & Engle, 1989) are really individual differences in a single component of the working memory system: controlled attention. Further, the differences in working memory capacity are not really differences in *memory* at all. Those differences reflect differential ability to use controlled attention to raise the activation of knowledge units, to maintain or sustain that activation in the face of interference and distraction, and, occasionally, to select among schemes contending for action on the basis of the strength of their activation levels.

Working Memory Capacity and the Frontal Lobes[12]

Research and theorizing about short-term and working memory have benefited considerably from neuropsychological research. As memory models become more complex, we must be mindful of whether the structures and processes proposed for the memory system fit with what is known about the brain.

There is growing evidence that connects the functions we have attributed to the central executive to structures in the frontal lobes, particularly the prefrontal cortex (Duncan, 1995; Duncan, Emslie, Williams, Johnson, & Freer, 1996; Goldman-Rakic, 1987; Kimberg & Farah, 1993; Pennington, 1994; Shallice & Burgess, 1993). Goldman-Rakic (1987), for example, has used a delayed response task in which monkeys are shown a food pellet being placed under one of two objects. Then a screen blocks the monkey's view of the two objects for a period of time, after which the monkey could have the food if it picked the correct object on the first trial. Normal monkeys have no difficulty in representing the correct object in memory over the delay and do well on the task. However, removal of parts of the prefrontal cortex (particularly Brodmann's area 46) leads to an inability on the part of the monkey to retain the information over the delay. This finding suggests that the prefrontal cortex is important to maintaining the temporary representation of the location of the hidden food.

Shallice and his colleagues (Shallice & Burgess, 1993) have argued that the frontal lobes are an important part of the circuitry of the supervisory attentional system, which we previously described. Evidence of this link shows that damage to the frontal lobes, particularly the prefrontal cortex, leads to difficulty in doing tasks that require sustained controlled attention (Duncan et al., 1996; Luria, Karpov, & Yarbuss, 1966; Weinberger, 1993). Duncan (1995) has also made the connection between sustained controlled attention, general intelligence, and the frontal lobes. He argues that frontal lobe damage leads to a substantial decline in general fluid intelligence and the ability to sustain controlled attention. This conclusion, of course, would follow from the research connecting the constructs of working memory capacity, central executive, and *gF* (Engle, Tuholski, Laughlin, & Conway, 1998; Kyllonen & Christal, 1990).

Although most of these studies used animals or patients with frontal lobe damage, there is at least suggestive evidence linking sustained attention to the frontal lobes in normal, non-brain-damaged individuals as well. For example, individuals with low WM capacity as defined by tasks such as reading span and operations span show similar (albeit less devastating) patterns of performance in comparison to frontal patients on a variety of cognitive tasks. Patients with frontal damage show decreased performance on the verbal fluency task which asks subjects to retrieve as many examples of a category as they can (Benton, 1968; Milner, 1964;

Pendleton Heaton, Lehamen, Hulihen, 1982). Rosen and Engle (1997a) have shown the same pattern for low working memory subjects. Frontal patients also tend to perseverate on a strategy even after it is no longer useful (Drewe, 1974; Luria, 1966; Milner, 1963, 1964). Tuholski & Engle (1997) showed that low working-memory subjects persist in using an ineffective mental model strategy longer than do high working-memory subjects. Finally, frontal patients have been shown, in comparison to normal subjects, to be more vulnerable to interference and less able to suppress irrelevant or inappropriate information (Dempster, 1992; Freedman & Cermak, 1986; Knight, 1995; Knight, Scabini, & Woods, 1989; Leng & Parkin, 1989; Longmore, Knight, Menlus, & Htope, 1988; Shimamura Gershberg, Jurica, Mangels, & Knight, 1992; Stuss, 1991). Similarly, low working-memory subjects have shown greater effects of proactive interference in a Brown-Peterson paradigm (Kane and Engle, 1997). Rosen and Engle (1997b) showed that in an *A-B*, *A-C*, *A-B* paired-associate procedure, high working-memory subjects suppressed the list-1 responses during the learning of list 2 but that low working-memory subjects did not. Further, low working-memory subjects made many more intrusions from earlier lists than did high working-memory subjects, suggesting that the former group had not suppressed the intruding items as well as had the latter group. In contrast, frontal patients perform normally on tasks that can be done under proceduralized or automatized processing (Bianchi, 1922; Frith, Friston, Liddle, & Frackowiak, 1991; Fuster, 1980, Penfield & Evans, 1935). Likewise, Conway and Engle (1994) showed that high and low working-memory subjects did not differ in a retrieval task in the absence of response competition.

In summary, there is growing speculation and evidence on the relationship between the central executive and the frontal lobes. It should be noted that Goldman-Rakic's work with monkeys speaks to the storage of the temporary representation itself. The other work cited is more directed at the use of controlled attention to maintain activation of a representation, to suppress interfering representations, or to choose between contending actions. This area of research is particularly vital right now, and we can expect many new findings and ideas on the neuropsychology of working memory, controlled attention, and general fluid intelligence in the near future.

Conclusion

We have briefly reviewed the history of the idea of a temporary memory store distinct from long-term memory starting with Hebb's (1949) physiological theory of the reverberatory trace. The psychological theories that followed were based on the idea that the temporary short-term trace behaved according to different laws than did the structural or long-term memory trace. Unfortunately, these theories were overly simplistic, and the characteristics attributed to short-term memory under the multi-store theories were probably a result of the particular tasks used rather than the inexorable nature of the temporary trace. Modern theories allow for multiple types of representation, each of which may be differentially sensitive to interference and loss over time. Further, controlled attention is important for the maintenance and/or suppression of the representation over time and for resolution of conflict between automatically activated action schemas. The newer theories not only are more content valid, because they show statistical relationships with higher-order or real-world cognitive functions; they also appear to be soundly based in brain science.

Acknowledgments

During the preparation of this chapter, the authors' work was supported by grants F49620-93-1-0336 from the Air Force Office of Scientific Research and RO1-HD-27490-01A1 from the National Institute of Child Health and Human Development. We thank Mike Kane for his helpful comments and criticisms.

Notes

1. Keeping with a useful convention adopted by Atkinson and Shiffrin (1968), we will use the term *short-term store* and the acronym *STS* to refer to the hypothesized temporary memory store, trace, or representation. The term *short-term memory*, or *STM*, will be used when talking about tasks and the phenomena observed from tasks that putatively reflect the underlying STS.

2. Despite the fact that Hebb's idea that the strength of the long-term trace was a function of time in the reverberatory trace has been discredited, the principle is still alive and kicking today in the form of Hebb's rule, which is an integral part of how learning occurs in neural net or connectionist models of cognition (Rumelhart, Hinton & McClellend, 1986).

3. *Ceiling effect* refers to the problem occurring when performance is so high that two conditions cannot be distinguished from each other. If the 100% ceiling did not limit them, one could be higher than the other. A similar problem occurs when performance in two conditions is nearly zero—that is, floor effect.

4. It is likely that modality of presentation is important to whether this generalization is correct. Auditory presentation leads to higher recency than visual presentation, and there is evidence (Cantor & Engle, 1989) that the recency found with auditory presentation is preattentive.

5. There was considerable debate early in this literature as to whether the code used in STM tasks was acoustic (i.e., sound based; Conrad & Hull, 1964) or articulatory (i.e., speech based; Wickelgren, 1966). The issue was never resolved; hence, we will use the neutral term *phonological*. It should be pointed out, however, that articulatory code seems to have won the war because the *articulatory loop* is such an important element of the Baddeley and Hitch (1974) model.

6. HM is more than likely the basis for the Mr. Short Term Memory character.

7. Since Brown (1958) and Peterson & Peterson (1959) both published work using a similar task, the task is typically referred to as the *Brown-Peterson task*.

8. In taking this view, Crowder adopted a position much like that of his mentor, Melton (1963).

9. See LaPointe & Engle (1990) for an example of how sampling with and without replacement in short-term memory studies can differentially affect the results.

10. The term *working memory* had been used earlier (e.g., Douglas, 1967) but, as is often the case in science, the Zeitgeist was not ready for the term until later.

11. There is considerable controversy over this issue, but the debate is beyond the scope of this paper. The reader is referred to papers by Waters and Caplan (1966) and Deaton, Gernsbacher, Robertson, and Miyake (1995).

12. There is an extensive neuropsychological literature on aspects of working memory other than central executive, particularly the phonological loop, but a complete coverage is beyond our scope.

References

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology, 6,* 451–474.

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

Anderson, J. R. (1983b). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22,* 261–295.

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* (Vol. 2, pp. 89–95). New York: Academic Press.

Baddeley, A. D. (1966). The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quarterly Journal of Experimental Psychology, 18,* 302–309.

Baddeley, A. D. (1986). *Working Memory.* Oxford, England: Clarendon Press.

Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology, 49A,* 5–28.

Baddeley, A. D., Bressi, S., Della Sala, S., Logie, R., & Spinnler, H. (1991). The decline of working memory in Alzheimer's disease: A longitudinal study. *Brain, 114,* 2521–2542.

Baddeley, A. D., Grant, S., Wight, E., & Thomson, N. (1975). Imagery and visual working memory. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and Performance* (Vol. 5, pp. 205–217). Hillsdale, NJ: Erlbaum.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). New York: Academic Press.

Baddeley, A. D., & Hitch, G. (1994). Developments in the concept of working memory. *Neuropsychology, 8,* 485–493.

Baddeley, A. D., Lewis, V. J., & Vallar, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology, 36,* 233–252.

Baddeley, A. D., & Lieberman, K. (1980). Spatial working memory. In R. Nickerson (Ed.), *Attention and Performance* (Vol. 7, pp. 521–539). Hillsdale, NJ: Erlbaum.

Baddeley, A. D., Logie, R., Bressi, S., Della Sala, S., & Spinnler, H. (1986). Dementia and working memory. *Quarterly Journal of Experimental Psychology, 38A,* 602–618.

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.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14,* 575–589.

Baddeley, A. D., & Warrington, E. K. (1970). Amnesia and the distinction between long-and short-term memory. *Journal of Verbal Learning and Verbal Behavior, 9,* 176–189.

Baddeley, A. D., & Wilson, B. (1985). Phonological coding and short-term memory in patients without speech. *Journal of Memory and Language, 24,* 490–502.

Baddeley, A. D., & Wilson, B. (1988). Frontal amnesia and the dysexecutive syndrome. *Brain and Cognition, 7,* 212–230.

Benton, A. L. (1968). Differential effects of frontal lobe disease. *Neuropsychologia, 6,* 53–60.

Benton, S. L., Kraft, R. G., Glover, J. A., & Plake, B. S. (1984). Cognitive capacity differences among writers. *Journal of Educational Psychology, 76,* 820–834.

Bianchi, L. (1922). *The mechanism of the brain and the function of the frontal lobes.* Edinburgh, Scotland: Livingstone.

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology, 6,* 173–189.

Broadbent, D. (1958). *Perception and Communication.* Oxford, England: Pergamon Press.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology, 10,* 12–21.

Butterworth, B., Campbell, R., & Howard, D. (1986). The uses of short-term memory: A case study. *Quarterly Journal of Experimental Psychology, 38A,* 705–738.

Cantor, J., & Engle, R. W. (1989). The influence of concurrent load on mouthed and vocalized modality effects. *Memory & Cognition, 17,* 701–711.

Cantor, J., & Engle, R. W. (1993). Working memory capacity as long-term memory activation: An individual differences approach. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 19,* 1101–1114.

Case, R. D. (1985). *Intellectual development: Birth to adulthood.* New York: Academic Press.

Case, R. D., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology, 33,* 386–404.

Conrad, R. (1957). Decay theory of immediate memory. *Nature, 179,* 831–832.

Conrad, R. (1964). Acoustic confusion in immediate memory. *British Journal of Psychology, 55,* 75–84.

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology, 55,* 429–432.

Conway, A. R. A., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General, 123,* 354–373.

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.

Cowan, N. (1995). *Attention and Memory: An Integrated Framework.* New York: Oxford University Press.

Cowan, N., Day, L., Saults, J. S., Keller, T. A., Johnson, T., & Flores, L. (1992). The role of verbal output time in the effects of word length on immediate memory. *Journal of Memory & Language, 31,* 1–17.

Cowan, N., Keller, T., Hulme, C., Roodenrys, S., McDougall, S., & Rack, J. (1994). Verbal memory span in children: Speech timing clues to the mechanisms underlying age and word length effects. *Journal of Memory & Language, 33,* 234–250.

Cowan, N., Wood, N. L., & Bourne, D. N. (1994). Reconfirmation of the short-term storage concept. *Psychological Science, 5,* 103–106.

Craik, F. I. M., & Jacoby, L. L. (1975). A process view of short-term retention. In F. Restle, R. C. Shiffrin, N. J. Castellan, H. Lindman, & D. Pisoni, (Eds.), *Cognitive Theory* (pp. 173–192). Hillsdale, NJ: Erlbaum.

Craik, F. I. M., & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671–684.

Craik, F. I. M., & Watkins, M. J. (1973). The role of rehearsal in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 12,* 599–607.

Crowder, R. G. (1983). The demise of short-term memory. *Acta Psychologica, 50,* 291–323.

Crowder, R. G. (1993). Short-term memory: Where do we stand? *Memory & Cognition, 21,* 142–145.

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

Daneman, M., & Carpenter, P. A. (1983). Individual differences in working-memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466.

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language, 25,* 1–18.

Deaton, J., Gernsbacher, M.A., Robertson, R., & Miyake, A. (1995) Working memory span and lexical ambiguity: problems with lexical access. Paper presented at Midwester Psychological Association, Chicago.

Dempster, F. N. (1981). Memory span: Sources of individual and developmental differences. *Psychological Bulletin, 89,* 63–100.

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review, 12,* 45–75.

Douglas, R. J. (1967). The hippocampus and behavior. *Psychological Bulletin, 67,* 416–442.

Drewe, E. A. (1974). The effect of type and area of brain lesion on Wisconsin Card Sort Test performance. *Cortex, 10,* 159–170.

Duncan, J. (1995). Attention, intelligence, and the frontal lobes. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences* (pp. 721–733). Cambridge, MA: MIT Press.

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology, 30,* 257–303.

Ellis, N. C., & Hennelley, R. A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculations in Welsh and English. *British Journal of Psychology, 71,* 43–52.

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 18,* 972–992.

Engle, R. W., Carullo, J. J., & Collins, K. W. (1991). Individual differences in the role of working memory in comprehension and following directions in children. *Journal of Educational Research, 84,* 253–262.

Engle, R. W., Conway, A. R. A., Tuholski, S. W., & Shisler, R. J. (1995). A resource account of inhibition. *Psychological Science, 6,* 122–125.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A., (in press). Working memory, short-term memory and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General.*

Farah, M. J. (1984). The neurological basis of mental imagery: A componential analysis, *Cognition, 18,* 245–272.

Farah, M. J., Hammond, K. M., Levine, D. N., & Calvanio, R. (1988). Visual and spatial mental imagery: Dissociable systems of representation. *Cognitive Psychology, 20,* 439–462.

Freedman, M., & Cermak, L. S. (1986). Semantic encoding defecits in frontal lobe disease and amnesia. *Brain and Cognition, 5,* 108–114.

Frith, C. D., Friston, K., Liddle, P. F., & Frackowiak, R. S. J. (1991). Willed action and the prefrontal cortex in man: A study with PET. *Proceedings of the Royal Society London B, 244,* 241–246.

Fuster, J. M. (1980). *The prefrontal cortex: Anatomy, physiology, and neuropsychology of the frontal lobe.* New York: Raven.

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language, 28,* 200–213.

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language, 29,* 336–360.

Glanzer, M., & Cunitz, A. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5,* 351–360.

Glanzer, M., & Razel, M. (1974). The size of the unit in short-term storage. *Journal of Verbal Learning and Verbal Behavior, 13,* 114–131.

Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In F. Plum (Ed.), *Handbook of physiology: Section 1. The nervous system: Vol. 5, Higher functions of the brain.* Bethesda, MD: American Physiological Society.

Gregg, V. H., & Gardiner, J. M. (1984). Phonological similarity and enhanced auditory recency in longer-term free recall. *Quarterly Journal of Experimental Psychology, 36A,* 13–27.

Hanley, J. R., Young, A. W., & Pearson, N. A. (1991). Impairment of the visuo-spatial sketch pad. *Quarterly Journal of Experimental Psychology, 43A,* 101–126.

Hebb, D. O. (1949). *Organization of Behavior.* New York: Wiley.

Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye (Ed.), *Brain mechanisms and learning* (pp. 37–46). London: Oxford University Press.

James, W. (1890). *The principles of psychology.* New York: Henry Holt.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122–149.

Kahneman, D. (1973). *Attention and Effort.* Englewood Cliffs, NJ: Prentice-Hall.

Kane, M., & Engle, R. W. (1998). *The role of working memory capacity in proactive interference.* Manuscript in preparation.

Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior, 1,* 153–161.

Kiewra, K. A., & Benton, S. L. (1988). The relationship between information processing ability and note taking. *Contemporary Educational Psychology, 13,* 33–44.

Kimberg, D. Y., & Farah, M. J. (1993). A unified account of cognitive impairments following frontal lobe damage: The role of working memory in complex, organized behavior. *Journal of Experimental Psychology: General, 122,* 411–428.

Kintsch, W., & Buschke, H. (1969). Homophones and synonyms in short-term memory. *Journal of Experimental Psychology, 80,* 403–407.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85,* 363–394.

Knight, R. T. (1995). Escape from linear time: Prefrontal cortex and conscious experience. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences* (pp. 1357–1371). Cambridge, MA: MIT Press.

Knight, R. T., Scabini, D., & Woods, D. L. (1989). Prefrontal cortex gating of auditory transmission in humans. *Brain Research, 504,* 338–342.

Kolb, B., & Whishaw, I. Q. (1990). *Fundamentals of human neuropsychology.* New York: W. H. Freeman.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity. *Intelligence, 14,* 389–433.

Kyllonen, P. C., & Stephens, D. L. (1990). Cognitive abilities as determinants of success in acquiring logic skill. *Learning and Individual Differences, 2,* 129–160.

LaPointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 16,* 1118–1133.

Leng, N. R. C., & Parkin, A. J. (1989). Aetiological variations in the amnesic syndrome: Comparisons using the Brown-Peterson task. *Cortex, 25,* 251–259.

Levy, B. A. (1971). The role of articulation in auditory and visual short-term memory. *Journal of Verbal Learning and Verbal Behavior, 10,* 123–132.

Lockhart, R. S., & Craik, F. I. M. (1990). Levels of processing: A retrospective commentary on a framework for memory research. *Canadian Journal of Psychology, 44,* 87–112.

Logie, R. H. (1986). Visuo-spatial processing in working memory. *Quarterly Journal of Experimental Psychology, 38A,* 229–247.

Longmore, B. E., Knight, R. G., Menkes, D. J., & Hope, A. (1988). The experimental investigation of a case of alcohol-induced frontal lobe atrophy. *Neuropsychology, 2,* 77–86.

Luria, A. R. (1966). *Higher cortical functions in man.* London: Tavistock.

Luria, A. R., Karpov, B. A., & Yarbuss, A. L. (1966). Disturbances of active visual perception with lesions of frontal lobes. *Cortex, 2,* 202–212.

McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychological Review, 39,* 352–370.

Melton, A. W. (1963). Implications for short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior, 2,* 1–21.

Melton, A. W., & Irwin, J. M. (1940). The influence of degree of interpolated learning on retroactive interference and the overt transfer of specfic responses. *American Journal of Psychology, 53,* 173–203.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97.

Milner, B. (1963). Effects of different brain lesions on card sorting. Archives of *Neurology, 9,* 90–100.

Milner, B. (1964). Some effects of frontal lobectomy in man. In J. M. Warren & K. Akert (Eds.), *The frontal granular cortex and behavior* (pp. 313–334). New York: McGraw-Hill.

Milner, B. (1966). Amnesia following operation on the temporal lobes. In C. W. M. Whitty & O. L. Zangwell (Eds.), *Amnesia* (pp. 109–133). London: Butterworths.

Murdock, B. B. (1961). The retention of individual items. *Journal of Experimental Psychology, 62,* 618–625.

Norman, D. A. (1969). *Memory and Attention.* New York: Wiley.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation* (Vol. 4, pp. 1–18). New York: Plenum Press.

Papagno, C., Valentine, T., & Baddeley, A. D. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language, 30,* 331–347.

Papagno, C., & Vallar, G. (1992). Phonological short-term memory and the learning of novel words: The effect of phonological similarity and item length. *Quarterly Journal of Experimental Psychology, 44A,* 47–67.

Pascal-Leone, J. (1970). A mathematical model for the transition rule in Piagets' developmental stages. *Acta psychologia, 63,* 301–345.

Pendleton, M. G., Heaton, R. K., Lehman, R. A., & Hulihan, D. (1982). Diagnostic utility of the Thurstone Word Fluency Test in neuropsychological evaluations. *Journal of Clinical Neuropsychology, 4,* 307–317.

Penfield, W., & Evans, J. (1935). The frontal lobe in man: A clinical study of maximum removals. *Brain, 58,* 115–133.

Pennington, B. F. (1994). The working memory function of the prefrontal cortices. In M. M. Haith, J. B. Bensen, R. J. Roberts, & B. F. Pennington (Eds.), *The development of future oriented processes* (pp. 243–285). Chicago: University of Chicago Press.

Peterson, L. R., & Peterson, M. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58,* 193–198.

Posner, M. I., & Peterson, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13,* 25–42.

Posner, M. I., & Snyder, C. R. (1975). Attention and cognitive control. In R. Solso (Ed.), *Information Processing and Cognition: The Loyola Symposium* (pp. 55–85). Potomac, MD: Erlbaum.

Postman, L., Stark, K., & Fraser, L. (1968). Temporal changes in interference. *Journal of Verbal Learning and Verbal Behavior, 7,* 672–694.

Reitman, J. S. (1971). Mechanisms of forgetting in short-term memory. *Cognitive Psychology, 2,* 185–195.

Reitman, J. S. (1974). Without surreptitious rehearsal: information in short-term memory decays. *Journal of Verbal Learning and Verbal Behavior, 13,* 365–377.

Reisberg, D., Rappaport, I., & O'Shaughnessy, M. (1984). Limits on working memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 10,* 203–221.

Roediger, H. L. (1990). Implicit memory: Retention without remembering. *American Psychologist, 45,* 1043–1056.

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval, *Journal of Experimental Psychology: General* 126, 211–227.

Rosen, V. M., & Engle, R. W. (1998). Working memory and suppression. (Manuscript submitted).

Rumelhart, D. E., Hinton, G., & McClellend, J. L. (1986). *A general framework for parallel distributed processing* (Vol. 1). Cambridge, MA: MIT Press.

Rundus, D., & Atkinson, R. C. (1970). Rehearsal processes in free recall: A procedure for direct observation. *Journal of Verbal Learning and Verbal Behavior, 9,* 99–105.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 501–518.

Schneider, W., & Detweiller, M. (1987). A connectionist/control architecture for working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 21, pp. 53–119). New York: Academic Press.

Shallice, T., & Burgess, P. W. (1993). Supervisory control of thought and action. In A. D. Baddeley and L. Weiskrantz (Eds.), *Attention: Selection, awareness and control: A tribute to Donald Broadbent* (pp. 171–187). Oxford: Oxford University Press.

Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory scores: A neuropsychological study. *Quarterly Journal of Experimental Psychology, 22,* 261–273.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: Pt. 2. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84,* 127–190.

Shimamura, A. P., Gershberg, F. B., Jurica, P. J., Mangels, J. A., & Knight, R. T. (1992). Intact implicit memory in patients with focal frontal lobe lesions. *Neuropsychology, 30,* 931–937.

Schneider, W., & Detweiler, M. (1987). A connectionist/control architecture for working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 21). New York: Academic Press.

Stuss, D. T. (1991). Self, awareness, and the frontal lobes: A neuropsychological perspective. In J. Strauss & G. R. Goethals (Eds.), *The Self: Interdisciplinary Approaches* (pp. 255–278). New York: Springer.

Tuholski, S. W., & Engle, R. W. *Individual differences in working memory capacity and the use of mental models.* Manuscript submitted for publication.

Tulving, E. (1966). Subjective organization and effects of repetition in multi-trial free-recall learning. *Journal of Verbal Learning and Verbal Behavior, 5,* 193–197.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28,* 127–154.

Warrington, E. K., Logue, V., & Pratt, R. T. C. (1971). The selective impairment of auditory verbal short-term memory. *Neuropsychologica, 9,* 377–387.

Warrington, E. K., & Weiskrantz, L. (1973). An analysis of short-term and long-term memory defects in man. In J. A. Deutsch (Ed.), *The physiological basis of memory.* New York: Academic Press.

Waters, G. S., Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to leading comprehension. *The Quarterly Journal of Experimental Psychology 1996, 49 A,* 51–79.

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review, 72,* 89–104.

Weinberger, D. R. (1993). A connectionist approach to the prefrontal cortex. *Journal of Neuropsychiatry, 5,* 241–253.

Wickelgren, W. A. (1966). Distinctive features and errors in STM for English consonants. *Journal of the Acoustical Society of America, 39,* 388–398.

Wilson, B. A., & Baddeley, A. D. (1993). Spontaneous recovery of impaired memory span: Does comprehension recover? *Cortex, 29,* 143–159.