This is a draft of a chapter published in M.L. Courage & N. Cowan (eds.) (2009), *The development of memory in infancy and childhood*. London: Psychology Press.

Development of Working Memory In Childhood

Nelson Cowan
University of Missouri, U.S.A.

Tracy Alloway
University of Durham, U.K.

Introduction

Working memory is the small amount of information that can be kept accessible or "kept in mind" in the service of ongoing cognitive activity. It is of key importance in language comprehension and all sorts of problem-solving. We discuss what working memory means, what is special about it, and how it might be understood within models of cognition and human information processing. Then we examine various types of specific processes within working memory that may undergo developmental improvement. Given limited space, we do not cover working memory in infancy, for which a different set of experimental techniques is needed (see Diamond, 1985; Oakes & Bauer, forthcoming; Rovee-Collier & Cuevas, this volume), or in non-human animals. Finally, we consider some practical consequences of working memory in cognitive disorders relevant to education.

What is Working Memory?

William James (1890, p. 403) said "Every one knows what attention is." Maybe so but, for a related concept, working memory, researchers disagree on what it is. Miyake and Shah (1999) asked the authors of various chapters in their book to define working memory and the definitions differed considerably. Some (e.g., Cowan, 1999) took a rather a priori approach and described it as the ensemble of mental mechanisms that allow information to be held in a temporarily accessible form to allow thinking. Others included in the definition the specific mechanisms or structures that they thought to be part of working memory, such as automatic storage mechanisms dedicated to a particular kind of information such as speech sounds, or a central attention mechanism that controls cognitive performance. One way to approximate a consensus is to use the definition from Wikipedia, an on-the-web encyclopedia allowing readers to edit the entry. On 30 August, 2006, the definition at the web site (http://en.wikipedia.org/wiki/Working_memory) was as follows. "In cognitive psychology, working memory is a theoretical framework that refers to structures and processes used for temporarily storing and manipulating information." The term dates back to a book by Miller, Galanter, and Pribram (1960) theorizing that, in carrying out a complex behavior (e.g., comprehending a written paragraph, completing a math problem, or packing a suitcase), an individual has to remember the overall goal memory as well as the data, an overall plan, and various sub-plans necessary to reach the goal. This information that is only temporarily appropriate to guide action is saved in what they called working memory.

Many other related terms have cropped up in the psychological literature, such as James' (1890) distinction between primary memory, comprising limited information in conscious awareness, and secondary memory, comprising the information learned over a lifetime. In the dawning of cognitive psychology in the late 1950's and 1960's, the terms short-term and long-term memory often served a similar purpose (e.g., Atkinson & Shiffrin, 1968). However, Baddeley and Hitch (1974) returned the field to the term working memory for several reasons. First, they wished to emphasize the use of this memory for carrying out cognitive operations; it is primarily from their influence that the Wikipedia definition of working memory includes not only holding information, but also manipulating it. Second, whereas primary or short-term memory was typically conceived as a single component of the mind, evidence by Baddeley and Hitch suggested that the concept had to be fractionated into several components, including separate phonological and visuospatial storage mechanisms that hold special information for a short time effortlessly, and a central processing unit that could hold various information while changing or manipulating it as the problem required (e.g., remembering numbers while adding them together). Third, they found that, although one could impede working memory by adding a demanding secondary task to be carried out along with the memory task, it did not affect some aspects of memory that other researchers attributed to primary memory; it did not quash the especially good recall of the most recent items in the list, or recency effect. Therefore, they came to see working memory as something different from what other people were calling primary or short-term memory. Nevertheless, all three terms -- primary, short-term, or working memory -- included a component of the mind that is very much like a focus of attention, a mindful, general-purpose entity that could deal with only a limited amount of information at once but could do wonders with it. Some have defined working memory as the active, attention-related part of temporary memory, and short-term memory as the passive part (e.g., Engle et al., 1999); others have defined working memory as the entire ensemble of mechanisms that includes both the active and passive parts (e.g., Baddeley, 1986; Cowan, 1988, 1999).

What is Special About Working Memory?

Ordinarily, what is special about a mental process is what it is capable of doing. However, to understand working memory, one must also consider what its limits are, as they distinguish it from the rest of the cognitive system. Working memory is capable of making certain ideas more readily accessible to ongoing thinking and actions (e.g., items, words, concepts, visual representations). However, it can do so only for a limited number of ideas at a time, and only for a limited time per idea. Therefore, there is an important process of selecting the ideas to be included in working memory. For example, if the goal of an action is lost from working memory, the action is not likely to be completed successfully until it is recovered. (Have you ever entered a room wondering, "Why did I come in here again?") If the part of the present sentence that precedes the comma is lost from working memory, then the part of the sentence after the comma will have no particular meaning for you.

Working Memory Limits

One can draw an analogy between the limits of working memory and limits of physical objects or events (cf. Kail & Salthouse, 1994). The latter are limited in time, space, and/or energy, which will be described in turn. Although these may encompass the basic working-memory mechanisms, there are other, more general processes that also influence performance on working memory tasks, and these too will be explained.

Time-related limits. In working memory, a time limit would indicate that any active idea in working memory stays active or accessible only for a short time, unless something is done to reactivate it, such as actively devoting attention to it anew or saying it to oneself (rehearsing it). This source of forgetting from working memory (or from short-term memory) is termed decay. In the early days of cognitive psychology, one could find evidence seeming to confirm decay (e.g., Baddeley, Thomson, & Buchanan, 1975; Peterson & Peterson, 1959) or seeming to disconfirm it (e.g., Keppel & Underwood, 1962; Waugh & Norman, 1965). Today, despite considerable additional research, one still can find recent evidence seeming to confirm decay (e.g., Barrouillet, Bernardin, & Camos, 2004; Cowan, Nugent, Elliott, & Geer, 2000; Mueller, Seymour, Kieras, & Meyer, 2003) or seeming to disconfirm it (e.g., Cowan et al., 2006; Lewandowsky, Duncan, & Brown, 2004; Lovatt, Avons, & Masterson, 2002). A challenge for this research is to be sure that rehearsal is not taking place. One can insert a distracting task to prevent rehearsal, but the distracting task may displace the memoranda in working memory, a mechanism different from decay. So, today, the proposition that there is forgetting of unrehearsed items in working memory as a function of time in the absence of interference (i.e., decay) is still quite controversial.

A corollary to the decay hypothesis is that the speed of processing should make a difference. For example, if items from a list are being forgotten while an individual is recalling them from working memory, then faster recall should allow more items to be recalled before they become unavailable for recall due to decay. Lewandowsky et al. (2004, Experiment 2) slowed down recall by having participants inserting a spoken word (super) once, twice, or thrice in each gap between letters from a list as it was recalled, but no effect of the amount of delay on recall was observed. Nevertheless, some would argue that when a single word is to be repeated like this, the repetition becomes routine enough that some attention is freed up and is used to reactivate the items to be recalled, which obscures the process of memory loss (Baddeley, 1986; Barrouillet et al., 2004).

Space-related limits. The relevance of space in working memory applies if one thinks of each idea as a physical object, in which case a metaphor for working memory is a box that holds the currently active objects. The number of objects that can be held depends on the space taken up by the box (i.e., its volume). This approach was advanced by one of the best-known articles in the field of psychology (Miller, 1956), in which it was suggested that adults seem to be able to remember lists of about seven items. According to a previous approach, it could have been expected that what is important for memorability of a list is the amount of information that each item carries, in addition to the number of items in the list. For example, there are 26 English letters so, to identify any letter, up to five binary (yes-or-no) questions must be answered (because $2^5 > 26 > 2^4$). If we learn that a letter is after m, it can be any of 13 letters; if is also before t, any of 6 letters; if also after p, any of 3 letters (q, r, or s); if also before s, either of 2 letters (q or r); and if it is also after q, it can only be r. There are only 10 digits (carrying less information per item than letters) and there are thousands of words (carrying much more information per item than letters); yet, we can remember lists of about 7 items no matter whether they are digits, letters, or words.

What increases this limit is grouping items together to form larger meaningful items, as in the acronym IBM, which includes three letters but forms a single meaningful unit representing a corporation in the United States (International Business Machines). Miller (1956) called this grouping process chunking. It is difficult to remember a list of 9 random letters, but it is easy when the letters form a much smaller number of meaningful chunks (as in the list IBM-NBC-RCA, representing 3 well-known corporations). Now, one could ask whether it is possible to remember lists of 7 of these acronyms, or, say, 7 idiomatic expressions or phrases, and the answer appears to be "no" (Glanzer & Razel, 1974; Simon, 1974; Tulving & Patkau, 1962). However, it has been argued that the ordinary limit of 7 is possible only because some special processes enhance working memory capability. Participants may group items together rapidly to form new chunks, or they may engage in silent, covert rehearsal of short verbal items (e.g., Baddeley, 1986). When those processes are prevented, participants more typically recall 3 or 4 chunks (Broadbent, 1975; Cowan, 2001) and this appears to be true not only for short chunks such as digits, letters, or words, which can be articulated quickly, but also for longer chunks such as phrases, or for chunks that cannot be articulated at all (e.g., Glanzer & Razel, 1974; Gobet & Clarkson, 2004).

Can these findings be accommodated with the space metaphor? Probably so. A box can hold only a fixed number of objects of a certain size, but the density of the objects does not matter. Chunking may be compared to taking objects of a standard size and compacting two or more of them together until they take up the same space that a single standard object took before the compression. By doing so, more objects can fit in the same box. Although there have been recent studies showing that capacity limits provide an excellent account of performance (Cowan, Chen, & Rouder, 2004; Chen & Cowan, 2005), the generality of this conclusion remains unknown.

Energy-related limits. The energy metaphor takes into account that information does not just enter working memory automatically and then leave it automatically. A person's goals make him or her strive to put certain items in working memory and to remove other, irrelevant items to make room for those relevant to the task at hand, whatever it may be. Yet, these goals can conflict with other mental tendencies that have to be suppressed or overridden. For example, in the classic procedure of Stroop (1935), the participant must name an ink color as quickly as possible instead of reading aloud the conflicting color word that the ink forms (e.g., the word blue written in yellow ink). Given that the conversion from print to speech is automatic in normal adults, the printed word must be inhibited to prevent errors in this task. Individuals with less working memory have a difficult time suppressing the word-reading response and can be more often lulled into reading the word instead of naming the ink color, especially if the word matches the ink color on most trials (Kane & Engle, 2003). Low-span individuals may have less of the relevant type of energy to control attention. In daily life, considerable energy must be exerted to maintain working memory, as when one waits patiently for a turn to say something without either interrupting or forgetting what one wanted to say. This general concept of working memory limits in energy (i.e., control of attention) may be less controversial than limits in time (decay) and space (capacity). However, in any particular situation, it is difficult to identify exactly which of these limits applies. For example, the inability to remove irrelevant items from working memory could deflate the estimate of how many relevant items can be held at once (May, Hasher, & Kane, 1999; Vogel, McCollough, & Machizawa, 2005).

Other limits affecting working memory. Last, a number of other factors may influence how well one performs a working memory task. Knowledge of the material being tested makes a difference (Chi, 1978), and its influence can be separated from those of other factors, such as recall speed (Hulme, Maughan, & Brown, 1991) or working memory for less familiar materials (Hambrick & Engle, 2001). The ability to use particular mnemonic strategies, such as types of rehearsal and grouping, makes a difference (Budd, Whitney, & Turley, 1995; McNamara & Scott, 2001), as does one's knowledge of what strategies would or would not be feasible or helpful, termed metamemory (Bunnell, Baken, Richards-Ward, 1999). With all of these possible factors, it is actually quite difficult to conclude that a particular age difference in memory definitely is caused by one particular factor.

With these ideas and cautions in mind, we now introduce some conceptions of the working memory system and then examine evidence of developmental change.

Models of Working Memory

Why Models are Helpful

It might be possible to test each proposition about working memory and its development separately. Is there decay? Is there a capacity limit? Is control of attention an issue? Which of these changes with development? In practice, though, it is difficult to do this. For example, if one wants to examine age differences in decay, one must explore the possibility that the combination of capacity and control-of-attention factors might produce results that look like decay. In exploring any one factor, it helps to keep in mind one's conception of the processing system as a whole. In any case, an important goal of research is to refine our understanding of the system as a whole. At least partly for these reasons, considerable effort has gone into the development of models of working memory. What is meant by this is a simplified conception of the parts of working memory and how they operate together, a bundle of hypotheses about working memory that all fit a logically consistent system leading to testable predictions.

Types of Models

One can distinguish at least four types of models of working memory. They can be classified as cognitive versus neural on one hand, and as conceptual versus computational on the other hand. In cognitive, conceptual models, the aim appears to be to take into consideration all that is known about working-memory performance and to describe a plausible system of mechanisms that could produce the data. The predictions based on this sort of model are primarily ordinal; performance is predicted to be higher in one particular condition than in another. In cognitive computational models, additional assumptions are made explicit so that the model can yield predictions about the mathematical form of results. This generally can be done only by adopting some assumptions that cannot presently be checked, and only by limiting the scope of evidence that can be taken into consideration. In neural models, statements are made about the neural activities underlying working memory performance. Neural models also can be of the either conceptual or the computational variety although, at present, to create a neural, computational model one generally has to go "far out on a limb" with unsubstantiated assumptions. We do not advocate the use of any one type of model alone; they are not mutually exclusive, and there is a need for all of them until such time as a grand, unifying, conceptually and computationally explicit model of performance on both the cognitive and neural levels can be developed and verified. For now, though, we emphasize one model type (below).

Cognitive, conceptual models. Figure 1 summarizes three models of this type. The first model is a rough sketch of the entire human information processing system offered by Broadbent (1958) in a footnote within his book. (The terminology is modified here, as Broadbent's original terms conflict with more recent usage.) Coming at the awakening of the field of cognitive psychology, Broadbent distinguished between two types of temporary memory: a sensory memory that contains traces from sensory experience in all modalities and channels at once, but not for long; and a central processing unit that can accept only a small fraction of the information present in sensory memory and is, in that way, limited in capacity. These ultimately contribute to a long-term memory of experience. According to Broadbent's conception, the higher-level stores could feed information back to the lower-level stores to help encode and interpret information within them. In most subsequent models there is a level of sensory memory that is not interpreted on the basis of categorical knowledge. Atkinson and Shiffrin (1968) refined this model to make it more computational and to give a more vivid account of control processes, such as covert rehearsal, that presumably are involved in cycling information from one phase of storage to another.

The middle panel shows the model of Baddeley (2000). An earlier version by Baddeley (1986) has been probably the most influential model in the field of working memory per se. The newer version shown in the figure differs from the earlier version only in that the former model did not include the box labeled the episodic buffer. The original model was designed to explain why a single capacity-limited store, such as that of Broadbent (1958), cannot account for all of the working memory results. The idea of a capacity-limited region of memory implies that keeping in mind one type of information impairs memory for another kind of information. In reality, though, Baddeley and colleagues found little interference between verbal materials, on one hand, and visual-spatial materials, on the other hand, so separate types of memory storage were advocated for these two types of materials. Although Baddeley and Hitch (1974) considered that there also might be a capacity-limited, general type of memory storage running across domains, Baddeley (1986) did away with that type of storage for the sake of parsimony and suggested that every type of memory of interest could rely on phonological and/or visuospatial memory, with central executive processes that manipulate and process information, controlling its flow, but do not store it per se. Phonological and visuospatial memory were said to be limited by decay, for reasons we will explore later. Baddeley (2000) reconsidered the need for a more general type of memory storage, for various reasons. For example, one must often temporarily remember the links between verbal and visual sources of information, or a new arrangement of abstract ideas that are neither phonological nor visuospatial in form. To account for these types of memory, Baddeley (2000) added the episodic buffer to the model.

The bottom panel of Figure 1 shows a model proposed by Cowan (1988) and elaborated in later writing (Cowan, 1995, 1999, 2001). It was suggested as a general processing model, taking into account the role of attention processes in storage more than did the model of Baddeley (1986). It can be contrasted to the other two models in several ways. Broadbent's (1958) model includes the notion of an attention filter limiting how much information can be transferred from the sensory store to the capacity-limited store at any moment; it is generally limited to a single stream of input. In contrast, Cowan (1988) considered that there must be some way for at least sensory information from multiple streams of input to be transferred. In evidence that Broadbent summarized, when individuals attend selectively to one of several channels (such as one ear's message in a dichotic listening situation with different messages to the two ears) they remember little or nothing of the material in the unattended channel. However, they do notice when an unattended message changes abruptly in its physical properties (such as a switch from a male voice to a female voice in the unattended message). They sometimes also notice their own names in the message to be ignored (Moray, 1959), although that can be explained on the basis that some individuals' attention wanders from the assigned channel, allowing attention to a channel that was supposed to be ignored (Conway, Cowan, & Bunting, 2001). Cowan's (1988) model postulates that all incoming stimulation has some influence on the processing system, activating parts of long-term memory. For unattended stimuli, this activated information may or may not be limited to the gross physical features of the stimuli. If these features change abruptly, the stimuli not only activate long-term memory features, but also recruit attention. This in turn results in a deeper, more complete analysis of the changed stimuli. The assumption is that a mnemonic representation of the environment is built up and that orienting of attention (Sokolov, 1963) occurs after an individual processes stimuli that produce a discrepancy from that representation. As the representation is modified to reflect the change, the orienting response dies down as the individual habituates to the new status quo.

Cowan's (1988) model also can be compared to Baddeley (2000). One difference is that Cowan's model is less modular. In place of the phonological and visuospatial stores, it depicts activated portions of long-term memory. The intent was not to deny a difference between phonological and visuospatial representations. However, at present it is not clear how memory should be divided up. Perhaps the phonological store also works for nonverbal sounds. Perhaps there is a representation that is auditory and spatial, or tactile and spatial. Therefore, Cowan simply stated the hypothesis that the amount of interference between one stimulus and another depends on the similarity in the features of the two stimuli; it was thought that positing a specific taxonomy of the varieties of activated memory, or the exact nature and number of buffers, would be premature.

In addition to activated elements of long-term memory, Cowan supposed that a subset of that activated memory is in the focus of attention. Later work suggested that the focus typically includes 3 to 5 meaningful chunks of information in normal adults (Cowan, 2001, 2005). It may also include new associations between elements. In fact, all items present in the focus of attention at once may become inter-associated. If one pays attention while two people are being introduced, for example, then the focus of attention may hold information about which name goes with which person. Whether all of this information is successfully transferred to long-term memory is a question that probably cannot yet be answered. The focus of attention functions in a manner that may be similar to Baddeley's (2000, 2001) episodic buffer, except that Baddeley has not taken a definitive stand on the attention requirements of that buffer. Overall, there is a lot of similarity between the revised Baddeley model and the Cowan model. According to one use of terminology, the activated memory can be called short-term memory and the focus of attention can be called working memory (Engle et al., 1999).

Other types of models. Other varieties of models aside from the cognitive, conceptual ones are important but have perhaps only just begun to make an impact on developmental research. Cognitive, computational models include many models specific to serial recall processes (for a summary see Farrell & Lewandowsky, 2004). A few models of short-term working memory more generally are conceptually quite similar to the Cowan (1999) model, but in a computationally explicit way (Anderson, Bothell, Lebiere, & Matessa, 1998; Davelaar, Goshen-Gottstein, Ashkenazi, Haarman, & Usher, 2005; Oberauer & Kliegl, in press). In neural, conceptual models of working memory, areas of the brain have begun to be identified with particular processes related to working memory, providing important clues to guide future investigations even though we do not yet know much about each of these brain areas. In particular, emphasis has been placed on frontal lobe systems as a basis of the control of attention (e.g., Cowan, 1995; Kane & Engle, 2002) as well as parietal lobe systems as an important part of the seat of attention, or its focus (Cowan, 1995, 1999). The intraparietal and intraocular sulci have been found to respond to visual arrays in a working memory task in a manner that mirrors performance. Todd and Marois (2004) found these areas to increase in activity as the visual working memory load increased, and to reach an asymptotic level of activity when the behaviorally-defined memory capacity was reached. Xu and Chun (2006) replicated this and showed that some of the sub-areas respond to the number of objects regardless of their complexity; but they also found related areas that are influenced by the level of complexity in the stimuli, which also does affect working memory capacity. This leads to the notion that there is a limit in how many objects can be represented (about 4 at most) and another limit in the total amount of complexity (with the capacity lowered for complex objects), all in keeping with behavioral results (Alvarez & Cavanagh, 2004). Finally, neural, computational models have suggested why working memory capacity is limited. The suggestion is basically that the multiple features of several items can be represented at the same time if the neural substrates of the features of a given object all fire in unison and, within a short, fixed cycle, all items have a turn. If more than a few objects enter working memory, there is an increasing danger that the features of different items will be confused with one another (Lisman & Idiart, 1995). For example, one may forget whether the items included a red square and a green triangle or a green square and a red triangle, if several other items also were included in the set.

In practice, the division between model types is not so clean because, for example, some conceptual models may appeal to neural-like principles without including many actual neural details, computational models may have a conceptual wing that is not modeled, broad conceptual models may have a computational component that is itself not very broad or comprehensive, and so on. All of the models may be viewed as theoretical tools to guide empirical investigations.

Distinguishing Between Models Based on Mechanisms

Although it is difficult to determine which model is correct, mechanisms help to distinguish between models. In some models, time-related decay of memory is postulated (e.g., Baddeley, 1986; Broadbent, 1958; Cowan, 1999) whereas, in other models, time-related decay is not part of the model at all (e.g., Lewandowsky et al., 2004; Waugh & Norman, 1965). According to some models there should be a general capacity that is shared between different types of materials (Baddeley, 2001; Broadbent, 1958; Cowan, 1999) but, according to other models, there is no such capacity because retention is accomplished entirely in domain-specific stores (e.g., Baddeley, 1986). Also, in some models (e.g., Engle et al., 1999; Kane & Engle, 2003) the ability to control attention might entirely explain individual differences in capacity whereas, according to other models, more emphasis is placed on working memory storage limits as being separate from these attention-control differences (Cowan, 1999, 2001). Addressing the fundamental mechanisms is probably an efficient way to guide the selection among models.

Development of Working Memory

No doubt working memory improves with development but it is difficult to determine what basic mechanisms account for this improvement with maturation. We will consider the role of mechanisms we have already introduced. Given that mechanisms help to distinguish between models of working memory, it seems likely that developmental evidence on the role of these mechanisms in turn can help guide the selection of models. Thus, the relation between cognitive psychology and cognitive development is a two-way street.

Time-related (Decay and Speed) Factors

The examination of time-related factors has involved experimental methods in which a list of items is presented and the task is to recall the list in its presented order (in developmental studies, orally or by pointing to items). Sometimes, fixed list lengths are used and, other times, the list length is variable and the question is the length of list that can be recalled without error (a memory span task).

Speed-memory correlation. Traditionally, investigations of decay and speed co-occur. According to the logic developed by Baddeley et al. in understanding the word length effect, individuals could remember lists of about as many words as they could pronounce in 2 seconds. The maximal speed of pronunciation of a few words was assumed to provide an estimate of the speed of covert rehearsal, perhaps not a bad assumption inasmuch as covert and overt varieties of speech do seem to proceed at comparable rates (Landauer, 1962). The data suggested that the speed of rehearsal could account for memory span. The assumption was that the phonological memory trace is lost if it is not refreshed soon enough through covert verbal rehearsal, in a repeating phonological loop. On the assumption that everyone has the same decay rate and loses an unrehearsed memory trace in about 2 seconds, individual differences in memory can be accounted for by differences in the speed of covert rehearsal. This logic was later extended to account for developmental differences in span, the finding being that younger children speak more slowly and recall commensurately less than older children or adults (Hulme, Thomson, Muir, & Lawrence, 1984; Hulme & Tordoff, 1989; Kail, 1992; Kail & Park, 1994). An example of this relationship is shown in Figure 2. Similar arguments had been made regarding the relation between other sorts of short-term memory and processing speed, as well (Case, Kurland, & Goldberg, 1982).

A problem with the phonological loop interpretation was that other research suggests that young (e.g., 4-year-old) children, whose data fit the nice linear relation between speech rate and memory span, cannot carry out useful rehearsal (Flavell, Beach, & Chinsky, 1966), suggesting that the rate of rehearsal cannot be a determining factor. A slight modification of the theory still could work. Whether the phonological memory trace lasts long enough to result in successful recall should depend not only on the speed of covert rehearsal, but on the speed of overt, spoken responses. Specifically, memory can be lost during the response period. If the first item in the response is articulated too slowly or if the pause after it is too slow, memory for the second and following items can fade during that time; and so on. Evidence does suggest that the overt response speed is of interest, in that (1) young children show effects of speech rate on memory span, such as a word length effect, only when there is a spoken response, whereas older children show these effects even when the response is nonverbal (Henry, 1991); and (2) the speed of spoken responses in memory span tasks does dramatically increase with age in childhood (Cowan et al., 1994, 1998).

Speed-memory path of causation? A critical question that still needs to be examined is whether the chain of causation suggested by the phonological loop model really holds. Recall the Lewandowsky et al. (2004) finding that, in adults, response speed in a serial recall task could be slowed without impairing span. Of course, their use of repetitions of a single word to slow down recall may have been inadequate to stop rehearsal during that delay (see Barrouillet et al., 2004). However, discouraging results also have been obtained in studies designed with the intent of speeding up children's recall, to see if span increases in the expected manner. Hulme and Muir (1985) tried to train children to learn to produce words at a faster rate to see whether that would result in an increase in memory span, presumably by allowing the children to rehearse more quickly. They were unable to train children reliably so it was not possible to test the hypothesis. Cowan et al. (2006, Experiment 2) took a slightly different approach, with second-grade children (most of whom were 8 years old) recalling lists of digits. Instead of trying to increase the maximum rate of rehearsal, this study concentrated on the usual rate of recall. A baseline digit span measurement was obtained in every child and then, in a second phase of the experiment, half of the children were simply instructed to speak as quickly as possible without making errors. In a third phase, they could again determine their own speed (but generally chose to remain fast). In a control group, no speed instruction was given in any phase. Each phase also included some post-span trials with lengths adjusted relative to span, to equalize the difficulty level. The results are shown in Figure 3. Even though children in the experimental group increased their speech rate to one faster than the usual adult rate (as measured in a first experiment), there was absolutely no improvement in memory in any part of the experiment.

Caution is still in order. It is possible that each person's memory response occurs at some sort of ideal rate for that individual. He or she might find through experience that slower speeds cause memory loss whereas faster speeds require so much effort that there is no benefit to memory, due to mechanisms outside of the phonological loop. It is also possible that second-grade children engage in some sort of rehearsal between items in their responses and that different results would be obtained if even younger children were instructed to recall quickly. However, an alternative account would state that the correlation between speech rate and memory span, or even between speed and memory more generally (Kail & Salthouse, 1994), occurs for some other reason. It could be that the memory requirement of saying things quickly plays a role. It also could be that both speed and memory are independent indices of some general neural efficiency and that neither one really influences the other. This issue is difficult to resolve but we look forward to interesting research to resolve it.

Decay differences with age. Another important question is whether an untested simplifying assumption is valid. According to the usual phonological-loop-based account of age differences, the rate of decay does not change with age but the rate of rehearsal and speech does change. It is difficult actually to measure decay (if, in fact, it exists) because it is difficult to separate decay from interference effects. In a working memory task, if one does not include a distracting stimulus during a retention interval, one cannot rule out the possibility that the participant rehearsed the stimuli. Yet, such distracting stimuli may cause interference. A couple of studies addressed this issue by engaging children in a silent computer game that involved silently looking for pictures with names that rhyme (Cowan, Nugent, Elliott, & Saults, 2000; Saults & Cowan, 1996). Most stimuli were to be ignored but, once in a while, the computer game would be replaced with a request to name the last spoken word (Saults & Cowan) or digit list (Cowan et al.). Attention and mnemonic strategies would be prevented both during the presentation of the auditory stimuli, so that spoken stimuli could not be memorized, and during a retention interval of 1, 5, or 10 seconds, so that spoken stimuli could not be rehearsed. If the silent computer game caused interference, we might expect that it could impair performance for digits in all serial positions within the study of Cowan et al. However, what actually was found, for lists adjusted to each individual's span, was a pronounced difference between the rates of decay in second- versus fifth-grade children at the final serial position and no difference at any previous position (see Figure 4). This suggests that there may be an age difference in decay of an uninterrupted sensory memory but that it applies only to the last item spoken. Each item may interrupt sensory memory for previous items so that only the last list item benefits much from that sort of memory (cf. Balota & Engle, 1981). For the sake of phonological loop explanations, this finding only suggests a small modification. Again, more research is needed.

Space-related (Capacity) Factors

Working memory capacity limits in terms of units, increasing in number with maturation, form the fundamental basis of cognitive development according to the neoPiagetian school of thought (Burtis, 1982; Case, 1972; Pascual-Leone, 1970, 2005; Weiss, 1995). This type of hypothesis applies across child development. There are still few studies of the development of the capacity of working memory in chunks, mainly because the field is still uncertain how these chunks should be identified. Cowan et al. (2005) examined several working memory tasks across the elementary school years and in adulthood, and several of the tasks were selected to minimize the role of mnemonic strategies so the role of capacity could be observed. In those tasks, the assumption was that each presented item was recalled as a separate chunk of information. One task was the array-comparison task of Luck and Vogel (1997), in which the presentation is generally too brief and arbitrary to allow much of a role of grouping or covert rehearsal (on the latter see Morey & Cowan, 2004, 2005). Another was a running memory span task with a very fast, 4-items-per-second presentation rate. A list of digits of an unpredictable length (12 - 20 items) was presented and then the participant was to recall as many items as possible from the end of the list. Because it was not possible to rehearse, the only way to accomplish the task was to wait until the list ended and then load as many items as possible from a passive sensory or phonological store into working memory. A third measure involved memory for spoken lists that were unattended at the time they were presented, as the participant was busy with a silent computer game, ignoring most lists until an occasional recall cue was presented just after a list ended. These measures suggested that the capacity increases with child development, from just over 2 items in the early elementary school years toward an adult capacity of about 4 items on average (generally ranging from 2 to 6 in adults). This is shown for one such measure, memory for spoken lists that were ignored during their presentation, in the solid lines of Figure 5. (The dashed lines show how much was added by attention at the time of presentation; a bit more in older participants.) It is not yet clear whether Cowan et al. have fair measures of capacity or whether the critical difference is the development of the ability to use attention to load working memory and use it at the time of recall (cf. Vogel et al., 2005). Therefore, it is not yet known whether capacity truly increases with age, or whether the apparent increase is a byproduct of an increase in the efficiency of loading and using of working memory.

Another type of capacity is not measured in meaningful chunks per se, but in phonological segments. Gathercole and Baddeley (1989) presented children with nonsense words of varying numbers of syllables (e.g., *prug, mulfost*) and found that children who could repeat longer nonwords had larger vocabularies. Nonword repetition at 4 years of age predicted vocabulary a year later (Gathercole, Willis, Emslie, & Baddeley, 1992), whereas the reverse was not true in that age range, although vocabulary became a predictor of nonword repetition at an older age. Later work showed that children who were able to repeat longer nonwords also went on to excel at reading, and the trend was especially strong when the least word-like nonwords were considered (Gathercole, 1995). A reasonable interpretation is that some children have a higher capacity for phonological information than others do. Still, it is not clear just how to think about this ability; is it a matter of longer or more detailed phonological representations in some children than in others, or less decay in some than in others? That is not clear, nor is the reason for developmental improvements in this task.

In sum, then, we actually know very little about whether capacity, narrowly defined as the number of chunks or other basic units that can be kept active at once, increases with age. We certainly know that working memory increases with age, but there is still a lot of work left to do to understand how basic capacity contributes to that improvement.

Energy-related (Attention Control) Factors

If one thing is clear from the last few decades of cognitive development research, it is that there is a developmental improvement in the ability to control attention. This has been demonstrated in a wide variety of behavioral research (e.g., for a summary see Bjorklund & Harnishfeger, 1990) and, increasingly, in neuroimaging research (e.g., for a summary see Posner, 2004). The frontal lobes of the brain have a long course of maturation throughout childhood, and their ability to help control attention (e.g., Kane & Engle, 2002) improves accordingly. Several central executive functions involved in what is commonly termed working memory are thought to depend on this control of attention, such as maintaining a goal (Kane, Bleckley, Conway, & Engle, 2001), updating working memory (Oberauer, 2005), switching attention from one task to another, and inhibiting irrelevant information (Friedman et al., 2006; May et al., 1999). This topic has been of great interest medically, given the epidemic of attention deficit disorder in children (Barkley, 2003). The brevity of our coverage of the control of attention in children occurs not for a lack of information, but because it would require a whole book. We will focus on one germane question: what specific effect does this development have on working memory performance?

The work leading to an understanding of the effects of attention on working memory in childhood has largely adapted procedures used to study individual differences in adults. Daneman and Carpenter (1980) developed a technique combining the storage of information with processing episodes so as to tie up both storage and processing, which, according to early views of working memory, shared resources (Baddeley & Hitch, 1974). The *sentence span* procedure is depicted on the left-hand side of Figure 6. A sentence is presented, and the participant is required to comprehend the sentence and then also retain the last word in the sentence for later recall. The measure of span is how many sentence-final words can be recalled in order while the sentences are adequately comprehended, as measured by a later test. Case et al. (1982) reported on a related procedure in which fields of dots are counted and the count for each field is then reported (*counting span*, middle of Figure 6). Later, Turner and Engle (1989) extended this procedure to cases in which the processing episode was arithmetic and a word to be recalled was presented after each problem (*operation span*, right side of Figure 6). These tasks all are sometimes referred to as *complex span* tasks for lack of a better term. (Operation span might have been a good general term but it has come to refer only to the arithmetic version.) Complex span tasks typically produce very high correlations with intellectual aptitude tasks, much higher than the digit span task that is part of most intelligence tests, and the complex span tasks together have become the gold standard for the measurement of working memory (Conway et al., 2005). Adults with higher working memory spans are better able to ignore an irrelevant channel in selective listening (Conway et al., 2001), avoid looking at a suddenly-appearing object when the assigned task is to look away from it (Kane et al., 2001), and avoid naming an irrelevant color word in the Stroop task (Kane & Engle, 2003).

There is controversy about why complex span correlates well with aptitudes. Is it because of the contribution of all central executive functions to intelligence, or only because of the contribution of the process of updating working memory, as Friedman et al. (2006) found? Cowan et al. (2005) found that complex span tasks correlated well with working memory tasks that did not include a dual task but still discouraged mnemonic strategies, such as running span with a fast presentation rate and two-array comparisons. Both predicted aptitudes well. Cowan et al. suggested that complex spans are effective because mnemonic strategies are impeded by the processing episodes, allowing the task to measure a core capacity (cf. Barrouillet et al., 2004).

Various researchers have stated different opinions as to why complex working memory span increases with age in childhood. Towse, Hitch, & Hutton (1998) carried out a counting span task in which the order of displays varied. They found that scores were lowest when the largest displays occurred at the beginning of the list rather than the end, presumably because, with the largest displays at the beginning, the counts that were to be recalled had to be maintained for a longer time while subsequent displays were presented. Similar results were obtained in operation and reading span tasks. However, for the age range tested (8-11 years), the delay manipulation hurt all age groups about equally. Therefore, this delay effect does not seem to be the basis of the development of working memory performance, which was robust under both short-processing-first or long-processing-first conditions.

According to Cowan et al. (2005), complex spans work by preventing mnemonic strategies. In adults, they provide a much higher correlation with aptitude tests than does a simple digit span, which freely allows grouping and rehearsal processes. However, in children too young to engage in these mnemonic processes, even a simple digit span task should work well and correlate highly with aptitudes. That is exactly what Cowan et al. found. Converging on their interpretation, other recent developmental studies have found that in children, as in adults, performance on complex span tasks correlates with aptitudes better when the span task is conducted with a time limit rather than according to the participant's chosen speed, which further takes up attention and thereby prevents mnemonic strategies from being sneaked into the task (Conlin, Gathercole, & Adams, 2005; Friedman & Miyake, 2004; Lépine, Barrouillet, & Camos, 2005).

If this last conjecture is correct, the implication is that the developmental difference observed in complex span tasks may not result from age differences in the *control* of attention as one might assume, but rather from age differences in the *scope* of attention, which may be synonymous with working memory capacity in chunks. Scope and control of attention seem to be partly independent and partly overlapping in terms of individual and developmental differences (Cowan, Fristoe, Elliott, Brunner, & Saults, in press). This is not to say that developmental differences in the control of attention are unimportant in development generally, but that their importance can be seen more clearly in working memory tasks for which the use of strategies is critical and can make use of that control of attention, not in complex span tasks in which strategies may be curbed. The contribution of strategies will be discussed shortly. The control of attention also can be seen to influence development in working memory tasks for which it is necessary to inhibit irrelevant information. For example, the detrimental effect of irrelevant speech on serial recall of a printed list is larger in younger children (Elliott, 2002).

When an individual, including a child, is off-task (e.g., Gathercole & Alloway, 2005) it is difficult to distinguish between a failure of capacity or the scope of attention, so that the instructions were forgotten, versus a failure of the control of attention, so that the individual was unable to exert the energy necessary to avoid distraction. We believe that, in principle, this question can be resolved but, in practice, it is still unresolved.

Unsworth and Engle (in press) argued that the control of attention is important in working memory tasks not only to hold information in primary or short-term memory, but also to guide the retrieval of information from secondary or long-term memory. This proposal includes the common assumption that no task is pure in terms of the processes tapped and that the tasks termed complex working memory span tasks often benefit from some role of memorization in the task, just as in a long-term memory task. Cowan et al. (2003) showed one way in which it may play a role. They measured the silent time preceding recall in several types of span task and found that these are much longer in sentence span tasks than in counting span or digit span tasks, and that the time difference is greater in younger children (Figure 7). One account of these results is that, in sentence spans, it is possible to try to reinstate sentences from long-term memory to provide a context for recalling the sentence-final words. That strategy takes a fair amount of time, especially in younger children. With Unsworth and Engle, we can speculate that this age difference may be the result of older children's and adults' more effective use of attention in long-term retrieval.

Other Factors

In addition to basic properties of time (speed and decay), space (capacity), and energy (attention control), remaining factors to be considered are quite important and include things that are not part of working memory strictly speaking but that, in practice, help performance considerably. These include *knowledge* of the stimuli to be remembered in working memory tasks and *strategies* for carrying out those tasks. Knowledge and strategies are covered at length by Bjorklund (this volume) but we will briefly consider their relation to the development of working memory.

Knowledge. Disentangling multiple processes in working memory requires caution. For example, suppose one presented to children of different ages both words and nonwords and found that the advantage of words over nonwords (there is one) was larger in older children. One might be tempted to conclude that this difference is a result of greater lexical knowledge in the older children. However, this result alone would not be sufficient to specify the mechanism involved. It could be that words are pronounced more quickly than nonwords and that this difference in pronunciation speed underlies the word-nonword difference and is greater in more mature individuals.

Hulme et al. (1991) showed how effects of lexical information and pronunciation speed can be dissociated. They presented lists of short, medium, and long words and nonwords; each list was homogeneous in both lexical status and length. Word length and lexical status operate in different ways, as Figure 8 shows. Word length affected recitation speed (which also was measured) and resulted in a linear relation between speed and span. Lexical status did not change the slope of that linear relation, but strongly affected the intercept. Roodenrys, Hulme, and Brown (1993) carried out a similar study developmentally. They found that there was a large speed difference between 6- and 10-year-olds; the older group's data points were higher up on the regression line relating speech rate to memory span, indicating both faster speaking capability and better span. However, older children also showed a slightly larger benefit with the presentation of words as opposed to nonwords, and this does indicate a contribution of children's growing lexical knowledge to memory span. Another example was described above for the use of linguistic knowledge to aid sentence span (Cowan et al., 2003). It is likely that many other examples exist in which knowledge is important for a particular working memory task.

Strategies. There is something special about strategies. For the other possible mechanisms of developmental change we have discussed, the proposal was one of a quantitative change: older children have less decay, faster processing, more capacity, better control of attention, and more knowledge. For strategies, we are in a different position because the proposed change is qualitative. Strategies begin in later childhood that apparently are not used at all in early childhood (cf. Bjorklund, this volume; Miller & Seier, 1994; Siegler, 1996).

Strategies that can apply to working memory tasks vary depending on the task. Two of the most common are verbal rehearsal (e.g., Flavell et al., 1966; Ornstein & Naus, 1978) and sorting items by some category such as meaning or appearance to assist recall (e.g., Schneider, Kron, Hunnerkopf, & Krajewski, 2004). One issue is whether the child produces a certain strategy and another issue is whether that strategy is used effectively. In a longitudinal study, Schneider et al. reported that children actually change from non-strategic to strategic performers rather quickly across tasks. A utilization deficiency seems uncommon but, when it appears, it seems to persist in the individual for long after strategies are first used, and correlates with poor working memory.

Two studies illustrate the great importance of strategies, at least in working memory tasks that allow strategy use. Cowan, Cartwright, Winterowd, and Sherk (1987) examined adults' serial order memory for lists of phonologically dissimilar spoken words, such as *fish, hat,* and *spoon*, as well as lists of phonologically similar words, such as *bat*, *cat,* and *rat*. The phonologically dissimilar words are recalled better from at least 4 years of age, but previous work showed that the magnitude of this phonological similarity effect increased markedly with age. A simple explanation is that, as older children and adults learn to use covert verbal rehearsal, this improves their serial order memory for dissimilar lists but does little for the similar lists, which are easily confused with one another during the rehearsal process. Verifying this theoretical account, Cowan et al. found that the pattern was changed in adults when they were required to whisper the alphabet while listening to the stimuli through headphones. This task can be assumed to suppress covert verbal rehearsal. The adults under articulatory suppression yielded results similar to 5-year-old children in both level and pattern.

A more elaborate study with a similar outcome was conducted by Cowan, Saults, & Morey, 2006). The question was how verbal-spatial associations are saved. In a task administered to third- and sixth-grade children and adults, an array of 3 to 7 line-drawing "houses" was presented. Then an equal number of names was presented, one at a time, with each name centered in a house and then disappearing, with 1-second presentations and 1 second inter-word intervals. Then a probe word was presented centrally and the task was to use the mouse to drag it to the house in which it had appeared. We considered that there are two ways to accomplish this task. One way is to store the verbal-spatial associations; that kind of storage of abstract information may be attention-demanding, however (Baddeley, 2000; Cowan, 1999). Another way would be to rehearse the list of names while, in parallel, retaining the visual path of houses assigned to those names. After the probe name was presented, the two types of information would be used together. For example, if the probe was the second name in the list, then the response is the second house in the spatial path. Thus, by exploiting something like what Baddeley (1986) termed passive phonological and visuospatial buffers together, it might be possible to exert less effort in this task.

We examined this issue by constructing two versions of the verbal-spatial mapping task. In a *1-to-1* version, each house was assigned to a single name. In an alternative, *uneven* version, this was not the case. At least one house was left empty, while at least one house was assigned to two, non-adjacent names in the list. By making the spatial path confusing, this uneven condition discourages the parallel use of phonological and visuospatial stores (see Figure 9). The finding was that third-grade children showed a significant advantage for the uneven condition, the sixth-grade children showed no advantage at all, and the adults showed a significant advantage for the 1-to-1 condition. This is to be expected if the third-grade children cannot rehearse and therefore must use the verbal-spatial associative process to retain the items; in the uneven condition, there are fewer houses per trial to remember. The adults, who can rehearse, use the parallel method that is easier to carry out in the 1-to-1 condition. Finally, the pattern of performance of adults who repeated the word *the* as an articulatory suppression task was nearly identical to that of third-grade children (see Figure 10). Presumably, suppression knocked out rehearsal and what was left over was similar to the mental apparatus that the third-grade children were able to bring to bear in this task. This study illustrates that developing strategies comprise a very important aspect of working memory development.

Development and Models of Working Memory

The developmental research we have reviewed seems to demonstrate change in multiple mechanisms, with no mechanism ruled out. Still, it can help to distinguish between models of working memory, in at least two ways. One way is by demonstrating developmental changes in memory stemming from a mechanism that should not be important according to the model. For example, given that only some models include decay of working memory representations as a function of time over a few seconds (see above), only those models can account for developmental differences in decay rates (e.g., Figure 4). This seems particularly important given how difficult it has been to observe decay conclusively in standard adult procedures. Another way development can help is by demonstrating a mechanism more clearly in children than in adults, given that the mechanism may be obscured by the sophisticated strategies that adults use. Thus, given that some models do not include a type of storage that cuts across specific domains, it would be hard for them to explain the pattern of results shown in Figure 10, by young children and by adults in a condition in which rehearsal is suppressed. As discussed above, that pattern suggests that verbal-spatial associations are saved in working memory, but only when a serial rehearsal strategy is not possible (in young children) or is difficult to apply (when the spatial path is convoluted). In all, the results of these studies seem consistent with the types of models offered by Baddeley (2000, 2001) and Cowan (1999, 2001), which allow for a general, cross-domain type of storage, more domain-specific interference mechanisms of storage (through either separate stores or separate types of interference effects), decay over time, and attention-demanding executive processes as possible sources of developmental change. Yet, these conclusions are still tentative.

Practical Implications

Research on the development of working memory, aside from clarifying a key part of normal cognitive development, may be helpful in education and medicine. Klingberg et al. (2005) have suggested that training with working memory tasks can improve a child's performance and adjustment across a wide range of tasks, although we believe that this finding must be further explored. Gathercole and Alloway (2004, 2005) have shown that poor classroom adjustment in children often is an effect of an undiscovered working memory deficit preventing the child from remembering and following instructions as well as carrying out assignments. Below, we provide further detail about some of the implications of working memory for cognitive impairments of childhood. Notice how the models and mechanisms of working memory we have discussed selectively come into play in these disorders, providing guidance to research on the causes of, and treatments for, these disorders.

Working memory and language disorders

There are strong links between short-term memory and language impairments. Tasks involving immediate recall of short sequences are among the best indicators of language deficits, such as specific language impairment (Gathercole & Baddeley, 1990). When asked to repeat nonwords such as 'woogalamic' and 'blonterstapping,' children with this impairment perform at a level two years behind their peers (Archibald & Gathercole, in press). The novelty of the phonological forms of nonwords prevents participants from using lexical knowledge to support recall as they could do in recalling words they know. As a result, this task appears to mimic how children learn vocabulary.

Some have suggested that, the greater the processing load in a task, the more children with language impairments will struggle. For example, Montgomery (2000) found that children with specific language impairment were more impaired on a condition in which they were to repeat a sequence of words in order of their physical size (e.g., hearing *nut, head, cow, thumb* and producing *nut, thumb, head, cow*), compared to simply recalling the words in any order. Yet, children with specific language impairment may perform at levels comparable to age-matched controls on visuospatial working memory measures (Archibald & Gathercole, in press). The working memory deficit in this group appears specific to the verbal domain.

Working memory and reading

Reading disabilities include difficulties in word recognition, spelling, and reading comprehension. It has often been suggested that phonological short-term or working memory deficits cause reading disabilities. The idea would be that it is difficult to read if one cannot remember the string of phonemes that result from sounding out the word. Oddly, though, this hypothesis has not proved to be true. Tasks that require memory of nonwords or short series of randomly arranged words have been used. Children with reading disabilities do show problems on such tasks, but the phonological memory at one age does not seem to predict reading performance at an older age (Wagner et al., 1997; Wagner & Muse, in press).

Other tasks also were examined in this sort of study, such as *phonological awareness* tasks in which the child must dissect a word into its parts, such as rhyming words or words with the same initial sounds, and the ability to name pictures rapidly. In a five-year longitudinal study of several hundred children who were followed from kindergarten through fourth grade, the measures were correlated but a key difference between them was that, at three different time periods, only phonological awareness skills predicted individual differences in reading words (Wagner, Torgesen, & Rashotte, 1999). The phonological awareness tasks probably require phonological memory, but they also require other skills that are essential for reading, such as letter-to-sound correspondences. Indeed, illiterate adults cannot do these tasks well (Morais, Bertelson, Luz, & Alegria, 1986) because the sounds of speech are quite entangled acoustically and rudimentary reading skills are therefore needed to understand that each syllable can be further decomposed into its sounds.

Complex working memory tasks that combine storage and processing are more predictive of reading. Children with reading disabilities show significant and marked decrements on such tasks (Siegel & Ryan, 1989). In typically developing children, scores on working memory tasks predict reading achievement independently of measures of phonological short-term memory (e.g., Swanson, 2003) or phonological awareness (e.g., Swanson & Beebe-Frankenberger, 2004). This special contribution of complex working memory tasks has been explained as the result of limited capacity for processing and storage together in children with reading disabilities (De Jong, 1998). It may take considerable working memory capacity to keep in mind at once the relevant speech sounds and concepts necessary for successfully identifying words and comprehending text.

Working memory and mathematics

There is also a close relationship between mathematical skills and working memory but this is mediated by the age of the child as well as the task. Verbal working memory plays a strong role in math skills in 7-year olds (Bull & Scerif, 2001; Gathercole & Pickering, 2000) and is also a reliable indicator of mathematical disabilities in the first year of formal schooling (Gersten, Jordan, & Flojo, 2005). However, once children reach adolescence, working memory is no longer significantly linked to mathematical skills (e.g., Reuhkala, 2001). One explanation for this change is that verbal working memory plays a crucial role for basic arithmetic (both to learn arithmetic facts and to retain relevant data such as carried digits) but that as children get older other factors, such as number knowledge and strategies, play a greater role (e.g., Thevenot & Oakhill, 2005). Low working memory scores have been found to be closely related to poor computational skills (Bull & Scerif, 2001; Geary, Hoard, & Hamson, 1999) and poor performance on arithmetic word problems (Swanson & Sachse-Lee, 2001).

Visuospatial memory is also closely linked with mathematical skills. It has been suggested that visuospatial memory functions as a mental blackboard, supporting number representation, such as place value and alignment in columns, in counting and arithmetic (D'Amico & Guarnera, 2005; Geary, 1990; McLean & Hitch, 1999). Specific associations have been found between visuospatial memory and encoding in problems presented visually (e.g., Logie et al. 1994), and in multi-digit operations (e.g., Heathcote, 1994). Visuospatial memory skills also uniquely predict performance in nonverbal problems, such as sums presented with blocks, in pre-school children (Rasmussen & Bisanz, 2005). In contrast, the role of verbal short-term memory is restricted to temporary number storage during mental calculation (Fürst & Hitch, 2000; Hechet, 2002), rather than general mathematical skills (McLean & Hitch, 1999; Reuhkala, 2001).

A key question regarding the relationship between working memory and learning disabilities is whether working memory is simply a proxy for IQ. There is evidence that working memory tasks measure something different from general intelligence tests (e.g., Cain, Oakhill, & Bryant, 2004; Siegel, 1988). While IQ tests measure knowledge that the child has already learned, working memory tasks is a pure measure of a child's learning potential. Thus, working memory skills are able to predict a child’s performance in both literacy and numeracy, even after a child’s general abilities have been taken into account (Gathercole, Alloway, Willis & Adams, 2006; Nation, Adams,Bowyer-Crane, & Snowling, 1999; Stothard & Hulme, 1992; for a review see Swanson & Saez, 2003).

Working memory and attention disorders

Could working memory impairments be explained by attention problems? The standard manual of clinical psychology, DSM-IV, describes Attention Deficit and Hyperactivity Disorders (ADHD) as including three subtypes: (1) hyperactive-impulsive behavior or ADHD-H, (2) inattention and distractibility or ADHD-I, and (3) a combination of these behaviors or ADHD-C. There is widespread agreement that individuals with ADHD-H have marked impairments in one particular aspect of central executive function, namely the ability to inhibit responses that have been overlearned, becoming what is known as prepotent (e.g., Barkley, 1997; Sonuga-Barker et al., 2002). An example is performance in the Stroop (1935) procedure, in which one must suppress a known color word in order to name the conflicting color of print in which it is written. Impairments of advanced executive functions such as organization and planning may be found in the later childhood years (Barkley, 2003; Brown, 2002). Crucially, though, children with ADHD-H do not consistently show impairments in working memory. In an influential review, Pennington & Ozonoff (1996) concluded that there was no evidence of general working memory deficits in such children, a claim that has been borne out also in more recent studies (e.g., Adams & Snowling, 2001; Kerns et al., 2001; Ruckridge & Tannock, 2002). There is, however, evidence that ADHD groups including ADHD-I may be impaired on tasks involving memory for visual or spatial information (e.g., Geurts et al., 2004; Mehta, Goodyer, & Sahakian, 2004).

Barkley (2003) has suggested that, in contrast to ADHD-H, pervasive working memory deficits are a hallmark of the other major subtype of ADHD, which include inattentivity and distractibility. Behaviors associated with working memory failures (such as having trouble remembering what s/he is doing) have been found to be significantly higher than normal in children with these subtypes of ADHD (Gioia et al., 2000). They also may be impaired in the coordination of dual tasks (Karetekin, 2004), which require both retention of information and executive function to coordinate the retained information.

Is it possible that working memory deficits underlie inattentive behavior in ADHD? A recent observational study of children with very poor verbal working memory skills but otherwise normal cognitive profiles shed light on this question (Gathercole, Lamont & Alloway, 2006). The children’s behavior was observed during normal classroom activities. The behavioral characteristics of this group appear to be very similar to those of children with ADHD-I.

However, three types of error were very common: forgetting complex instructions, failing to cope with tasks that impose significant processing and storage loads, and losing their place in complex tasks. Interestingly, the teachers failed to identify the children as having poor memory skills, attributing their poor classroom performance instead to either lack of attention or poor motivation. There is not yet enough evidence to determine whether children diagnosed with inattentivity and those with impairments of verbal working memory represent a single group or separate groups. Based on the available data we speculate that the ADHD-I and memory-impaired groups both have deficits of verbal working memory, whereas children with ADHD-H are deficient in inhibitory processing and, possibly, planning and visuospatial working memory, but not in verbal working memory. More work is needed on the group with combined symptoms, ADHD-C.

We have shown that children with various deficits learning and cognitive performance show different types of working memory impairment; they can be impaired on verbal retention, spatial retention, and/or executive function, mirroring three main parts of the working memory model of Baddeley (1986). However, the evidence on children with working memory impairment is not yet sufficient to indicate which theoretical model of working memory is correct. In the model of Cowan (1988, 1999), for example, the activated portions of long-term memory are allowed to be divisible into different modalities and codes. Such divisions include, but are not necessarily limited to, the division between verbal and visuospatial information (cf. Alloway, Gathercole, & Pickering, 2006) . Theoretically, for example, it might be possible to find an individual with an inability to retain the spatial arrangement of sounds. It seems likely that there are also deficits related to other aspects of working memory, including improper functioning of the focus of attention (Cowan, 1988, 2005) or episodic buffer (Baddeley, 2000, 2001), which may result in abnormal development in the ability to retain associations between features of an object (e.g., Cowan, Naveh-Benjamin, Kilb, & Saults, in press). It is not clear if the latter deficits are part of ADHD. Developmental deficits in specific mechanisms such as decay, speed, and chunk capacity warrant further research.

Concluding Observations

Working memory is central to all of cognition. It is definite that working memory improves with age in childhood but that is not a very specific statement. We have considered possible developmental changes in decay, speed, chunk capacity, control of attention, knowledge, and mnemonic strategies. It remains to be determined whether the number of causal concepts can be reduced in further work; perhaps, for example, age differences in speed are a result, rather than a cause, of age difference in capacity. Strategic change is probably the strongest example of developmental change, reflecting an actual qualitative change in the nature of mnemonic processing rather than just a quantitative change. Still, we do not yet know if that type of change is the most fundamental or results from sufficient resources to allow the strategy to be carried out, as a neoPiagetian approach might suggest. It is clear that working memory is important not only theoretically, for an understanding of the development of the mind; but also for the practical understanding of developmental and educational disorders. When we understand what is fundamental in working memory, we will better understand how to treat the disorders, too.

References

Adams, J.W., & Snowling, M.J. (2001). Executive function and reading impairments in children reported by their teachers as 'hyperactive'. *British Journal of Developmental Psychology, 19*, 293-306.

Alloway, T.P., Gathercole, S.E., & Pickering, S.J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separable? *Child Development, 77,* 1698-1716.

Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A. M. (2005). Working memory and special educational needs. *Educational and Child Psychology, 22*, 56-67.

Alvarez, G.A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15*, 106-111.

Anderson, J.R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language, 38*, 341-380.

Archibald, L.M., & Gathercole, S. E. (in press). Short-term and working memory in Specific Language Impairment. *International Journal of Communication Disorders*.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89-195). New York: Academic Press.

Baddeley, A.D. (1986). *Working memory*. Oxford Psychology Series #11. Oxford: Clarendon Press.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in cognitive sciences, 4*, 417-423.

Baddeley, A. (2001). The magic number and the episodic buffer. *Behavioral and Brain Sciences, 24*, 117-118.

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105, 158-173.*

Baddeley, A.D., & Hitch, G. (1974). Working memory. In G. Bower (Ed.), *The Psychology of Learning and Motivation, 8, pp. 47-90*. New York: Academic Press.

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.

Balota, D. A., & Engle, R. W. (1981). Structural and strategic factors in the stimulus suffix effect. *Journal of Verbal Learning and Verbal Behavior, 20*, 346-357.

Barkley, R.A. (1997). *ADHD and the nature of self-control*. New York: Guilford.

Barkley, R.A. (2003). Issues in the diagnosis of attention-deficit/hyperactivity disorder in children. *Brain & Development, 25,* 77-83.

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General, 133,* 83-100.

Bjorklund, D.F., & Harnishfeger, K.K. (1990). The resources construct in cognitive development: Diverse sources of evidence and a theory of inefficient inhibition. *Developmental Review, 10,* 48-71.

Broadbent, D.E. (1958). *Perception and communication.* New York: Pergamon Press.

Broadbent, D.E. (1975). The magic number seven after fifteen years. In A. Kennedy & A. Wilkes (eds.), *Studies in long-term memory*. Wiley. (pp. 3-18)

Brown, T.E. (2002). DSM-IV: ADHD and executive function impairment. In *Proceedings of the Annual Meeting of the American Psychiatric Association, 2,* 910-914.
Budd, D., Whitney, P., & Turley, K.J. (1995). Individual differences in working memory strategies for reading expository text. *Memory & Cognition, 23,* 735-748.
Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability. Shifting, inhibition and working memory. *Developmental Neuropsychology, 19,* 273-293.
Bunnell, J.K., Baken, D.M., Richards-Ward, L.A. (1999). The effect of age on metamemory for working memory. *New Zealand Journal of Psychology, 28,* 23-29.
Burtis, P.J. (1982). Capacity increase and chunking in the development of short-term memory. *Journal of Experimental Child Psychology, 34*, 387-413.
Butterworth, G. (1977). Object disappearance and error in Piaget's Stage IV task. *Journal of Experimental Child Psychology, 23*, 391-401.
Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: concurrent prediction by working memory, verbal ability and component skills. *Journal of Educational Psychology, 96*, 31-42.
Case, R. (1972). Validation of a neo-Piagetian mental capacity construct. *Journal of Experimental Child Psychology*, *14*, 287-302.
Case, R., Kurland, D.M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology, 33*, 386-404.
Chen, Z., & Cowan, N. (2005). Chunk limits and length limits in immediate recall: A reconciliation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1235-1249.
Chi, M.T.H. (1978). Knowledge structures and memory development. In R. Siegler (Ed.), *Children's thinking: What develops*? Hillsdale, NJ: Erlbaum.
Conlin, J. A., Gathercole, S. E., & Adams, J. W. (2005). Children's working memory: Investigating performance limitations in complex span tasks. *Journal of Experimental Child Psychology, 90*, 303-317.
Conners, K. (1997). Conners' Teacher Rating Scale - Revised (S). New York: Multi-Health Systems Inc.
Conway, A.R.A., Cowan, N., & Bunting, M.F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin & Review, 8,* 331-335.
Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review & user's guide. *Psychonomic Bulletin & Review, 12*, 769-786.
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.* Oxford Psychology Series, No. 26. New York: Oxford University Press.
Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (eds.), *Models of Working Memory: Mechanisms of active maintenance and executive control*. Cambridge, U.K.: Cambridge University Press. (pp. 62-101)
Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87-185.
Cowan, N. (2005). *Working memory capacity*. Hove, East Sussex, UK: Psychology Press.

Cowan, N., Cartwright, C., Winterowd, C., & Sherk, M. (1987). An adult model of preschool children's speech memory. *Memory and Cognition*, *15*, 511-517.

Cowan, N., Chen, Z., & Rouder, J.N. (2004). Constant capacity in an immediate serial-recall task: A logical sequel to Miller (1956). *Psychological Science*, *15*, 634-640.

Cowan, N., Elliott, E.M., Saults, J.S., Morey, C.C., Mattox, S., Hismjatullina, A., & Conway, A.R.A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*, 42-100.

Cowan, N., Elliott, E.M., Saults, J.S., Nugent, L.D., Bomb, P., & Hismjatullina, A. (2006). Rethinking speed theories of cognitive development: Increasing the rate of recall without affecting accuracy. *Psychological Science*, *17*, 67-73.

Cowan, N., Fristoe, N.M., Elliott, E.M., Brunner, R.P., & Saults, J.S. (in press). Scope of attention, control of attention, and intelligence in children and adults. *Memory & Cognition*.

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 and Language, 33*, 234-250.

Cowan, N., Naveh-Benjamin, M., Kilb, A., & Saults, J.S. (in press). Life-Span development of visual working memory: When is feature binding difficult? *Developmental Psychology*.

Cowan, N., Nugent, L.D., Elliott, E.M., and Geer, T. (2000). Is there a temporal basis of the word length effect? A response to Service (1998). *Quarterly Journal of Experimental Psychology, 53A*, 647-660.

Cowan, N., Nugent, L.D., Elliott, E.M., & Saults, J.S. (2000). Persistence of memory for ignored lists of digits: Areas of developmental constancy and change. *Journal of Experimental Child Psychology, 76*, 151-172.

Cowan, N., Saults, J.S., & Morey, C.C. (2006). Development of working memory for verbal-spatial associations. *Journal of Memory and Language*, *55*, 274-289.

Cowan, N., Towse, J.N., Hamilton, Z., Saults, J.S., Elliott, E.M., Lacey, J.F., Moreno, M.V., & Hitch, G.J. (2003). Children’s working-memory processes: A response-timing analysis. *Journal of Experimental Psychology: General*, *132*, 113-132.

Cowan, N., Wood, N.L., Wood, P.K., Keller, T.A., Nugent, L.D., & Keller, C.V. (1998) . Two separate verbal processing rates contributing to short-term memory span. *Journal of Experimental Psychology: General*, *127*, 141-160.

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

Davelaar, E.J., Goshen-Gottstein, Y., Ashkenazi, A., Haarman, H.J., & Usher, M. (2005). The demise of short-term memory revisited: Empirical and computational investigations of recency effects. *Psychological Review*, *112*, 3-42.

De Jong, P. F. (1998). Working memory deficits of reading disabled children. *Journal of Experimental Child Psychology*, **70**, 75-96.

Diamond, A. (1985). Development of the ability to use recall to guide action, as indicated by infants’ performance on AB. *Child Development, 56*, 868 – 883.

Elliott, E.M. (2002). The irrelevant-speech effect and children: Theoretical implications of developmental change. *Memory & Cognition, 30*, 478-487.

Engle, R.W., Tuholski, S.W., Laughlin, J.E., & Conway, A.R.A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128*, 309-331.

Farrell, S., & Lewandowsky, S. (2004). Modelling transposition latencies: Constraints for theories of serial order memory. *Journal of Memory and Language, 51*, 115-135.

Feigenson, L., & Carey, S. (2002). Infants' Discrimination of number vs. continuous extent. *Cognitive Psychology, 44*, 33–66.

Feigenson, L. & Carey, S. (2003). Tracking individuals via object-files: Evidence from infants' manual search. *Developmental Science*, *6*, 568-584.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*, 307-314.

Feigenson, L., & Halberda, (2004). Infants chunk object arrays into sets of individuals. *Cognition*, *91*, 173-190.

Flavell, J.H., Beach, D.H., & Chinsky, J.M. (1966). Spontaneous verbal rehearsal in a memory task as a function of age. *Child Development, 37*, 283-299.

Friedman, N. P., & Miyake, A. (2004). The reading span test and its predictive power for reading comprehension ability. *Journal of Memory and Language, 51,* 136–158.

Friedman, N.P., Miyake, A., Corley, R.P., Young, S.E., DeFries, J.C., & Hewitt, J.K. (2006). Not all executive functions are related to intelligence. *Psychological Science, 17*, 172-179.

Fürst, A. J., & Hitch, G. J. (2000). Different roles for executive and phonological components of working memory in mental arithmetic. *Memory & Cognition, 28*, 774–782.

Gathercole, S.E. (1995). Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. *Memory & Cognition*, *23*, 83-94.

Gathercole, S. E. (2004). Working memory and learning during the school years. *Proceedings of the British Academy*, *125*, 365-380.

Gathercole, S.E, & Alloway, T.P. (2004). Working memory and classroom learning. *Dyslexia Review*, *15*, 4-9.

Gathercole, S.E. & Alloway, T.P. (2005).*Understanding working memory: A classroom guide.* Medical Research Council.

Gathercole, S.E., Alloway, T.P., Willis, C., & Adams, A.M. (2006). Working memory in children with reading disabilities. *Journal of Experimental Child Psychology, 93,* 265-281.

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

Gathercole, S.E., Lamont, E., & Alloway, T.P. (2006). Working memory in the classroom. In S. Pickering (Ed.). *Working memory and education*, pp. 219 -240. Elsevier Press

Gathercole, S. E. & Pickering, S. J. (2000). Assessment of working memory in six- and seven-year old children. *Journal of Educational Psychology*, *92*, 377-390.

Gathercole, S.E., Willis, C., Emslie, H., & Baddeley, A. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology, 28*, 887-898.

Geary. D. C. (1990). A componential analysis of an early learning deficit in mathematics. *Journal of Experimental Child Psychology, 49,* 363-383.

Geary, D.C., Hoard, M.K., & Hamson, C.O. (1999). Numerical and arithmetical cognition: Patterns of functions and deficits in children at risk for a mathematical disability. *Journal of Experimental Child Psychology, 74,* 213-239.

Gersten, R., Jordan, N.C., & Flojo, J.R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities*, *38*, 293-304.

Geurts, H.M., Vertie, S., Oosterlaan, J., Roeyers, H., Sergeant, J.A. (2004). How specific are executive functioning deficits in attention deficit hyperactivity disorder and autism? *Journal of Child Psychology and Psychiatry, 45*, 836-854

Gioia, Isquith, Guy, & Kenworthy, L. (2000). *Behavior Rating Inventory of Executive Function*. London: Harcourt Assessment.
Glanzer, M., & Razel, M. (1974). The size of the unit in short-term storage. *Journal of Verbal Learning & Verbal Behavior, 13*, 114-131.
Gobet, F., & Clarkson, G. (2004). Chunks in expert memory: Evidence for the magical number four… or is it two? *Memory, 12*, 732-747.
Hambrick, D.Z., & Engle, R.W. (2001). Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. Cognitive Psychology, 44, 339-387.
Heathcote, D. (1994). The role of visuospatial working memory in the mental addition of multi-digit addends. *Current Psychology of Cognition*, 13, 207-245.
Henry, L.A. (1991). The effects of word length and phonemic similarity in young children's short-term memory. *Quarterly Journal of Experimental Psychology, 43A*, 35-52.
Hulme, C., Maughan,S., & Brown, G.D.A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory & Language*, *30*, 685-701.
Hulme, C., & Muir, C. (1985). Developmental changes in speech rate and memory span: A causal relationship? *British Journal of Developmental Psychology, 3*, 175-181.
Hulme, C., Thomson, N., Muir, C., & Lawrence, A. (1984). Speech rate and the development of short-term memory span. *Journal of Experimental Child Psychology, 38*, 241-253.
Hulme, C., & Tordoff, V. (1989). Working memory development: The effects of speech rate, word length, and acoustic similarity on serial recall. *Journal of Experimental Child Psychology, 47*, 72-87.
James, W. (1890). *The principles of psychology*. NY: Henry Holt.
Kail, R. (1992). Processing speed, speech rate, and memory. *Developmental Psychology, 28*, 899-904.
Kail, R., & Park, Y.-S. (1994). Processing time, articulation time, and memory span. *Journal of Experimental Child Psychology, 57*, 281-291.
Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica, 86*, 199-255.
Kane, M.J., Bleckley, M.K., Conway, A.R.A., & Engle, R.W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*, 169-183.
Kane, M.J., & Engle, R.W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review, 9*, 637-671.
Kane, M.J., & Engle, R.W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132*, 47-70.
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.
Kerns, K.A., McInerney, R.J., & Wilde, N.J. (2001). Time reproduction, working memory, and behavioural inhibition in children with ADHD. *Child Neuropsychology, 7*, 21-31.
Klingberg, T., Fernell, E., Olesen, P.J., Johnson, M., Gustafsson, P., Dahlstrom, K., et al. (2005). Computerized training of working memory in children with ADHD – A randomized,

controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*, 177-186.

Landauer, T.K. (1962). Rate of implicit speech. *Perceptual & Motor Skills, 15*, 646.

Lépine, R., Barrouillet, P., & Camos, V. (2005). What makes working memory spans so predictive of high level cognition ? *Psychonomic Bulletin & Review, 12*, 165-170.

Lewandowsky, S., Duncan, M., & Brown, G.D.A. (2004). Time does not cause forgetting in short-term serial recall. *Psychonomic Bulletin & Review, 11*, 771-790.

Lisman, J.E., & Idiart, M.A.P. (1995). Storage of 7 + 2 short-term memories in oscillatory subcycles. *Science, 267*, 1512-1515.

Logie, R.H., Gilhooly, K.J., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory & Cognition*, 22, 395-410.

Lovatt, P., Avons, S.E., & Masterson, J. (2002). Output decay in immediate serial recall: Speech time revisited. *Journal of Memory and Language, 46*, 227-243.

Luck, S.J., & Vogel, E.K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279-281.

May, C.P., Hasher, L., & Kane, M.J. (1999). The role of interference in memory span. Memory & Cognition, 27, 759-767.

McLean, J.F., & Hitch, G.H. (1999). Working memory impairments in children with specific mathematics learning difficulties. *Journal of Experimental Child Psychology*, 74, 240-260.

McNamara, D.S., & Scott, J.L. (2001). Working memory capacity and strategy use. *Memory & Cognition*, *29*, 10-17.

Mehta, M.A., Goodyer, I.M., Sahakian, B.J. (2004). Methylphenidate improves working memory and set-shifting in AD/HD: relationships to baseline memory capacity. *Journal of Child Psychology and Psychiatry, 45*, 292-305.

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.

Miller, G.A., Galanter, E, and Pribram, K.H. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart and Winston, Inc.

Miller, P. H., & Seier, W. L. (1994). Strategy utilization deficiencies in children: When, where, and why. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 25, pp. 107–156). San Diego: Academic Press.

Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in man. *British Medical Bulletin, 27*, 272–277.

Miyake, A., & Shah, P. (eds.) (1999). *Models of Working Memory: Mechanisms of active maintenance and executive control*. Cambridge, U.K.: Cambridge University Press.

Montgomery, J. (2000). Verbal working memory in sentence comprehension in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *43*, 293-308.

Morais, J., Bertelson, P., Luz, C., & Alegria, J. (1986). Literacy training and speech segmentation. *Cognition, 24*, 45-64.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology, 11*, 56-60.

Morey, C.C., & Cowan, N. (2004). When visual and verbal memories compete: Evidence of cross-domain limits in working memory. *Psychonomic Bulletin & Review*, *11*, 296-301.

Morey, C.C., & Cowan, N. (2005). When do visual and verbal memories conflict? The importance of working-memory load and retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 703-713.

Mueller, S.T., Seymour, T.L., Kieras, D.E., & Meyer, D.E. (2003). Theoretical implications of articulatory duration, phonological similarity, and phonological complexity in verbal working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 6*, 1353–1380.

Nation, K., Adams, J.W., Bowyer-Crane, C.A., & Snowling, M.J. (1999). *Journal of Experimental Child Psychology, 73*, 139-158.

Oakes, L.M., & Bauer, P.J. (eds.), (forthcoming), *Short- and long- term memory in infancy and early childhood*. New York: Oxford University Press.

Oberauer, K. (2005). Control of the contents of working memory—A comparison of two paradigms and two age groups. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 714–728.

Oberauer, K. & Kliegl, R. (in press). A formal model of capacity limits in working memory. *Journal of Memory and Language*.

Ornstein, P.A., & Naus, M.J. (1978). Rehearsal processes in children's memory. In P.A. Ornstein (ed.), *Memory development in children*. Hillsdale, NJ: Erlbaum. (pp. 69-99)

Pascual-Leone, J. A. (1970). Mathematical model for the transition rule in Piaget's developmental stages. *Acta Psychologica, 32,* 301-345.

Pascual-Leone, J. (2005). A neoPiagetian view of developmental intelligence. In O. Wilhelm & R.W. Engle (Eds.), *Understanding and measuring intelligence*. London: Sage. (pp. 177-201)

Pennington, B.F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry, 37*, 51-87.

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

Posner, M.I. (Ed.). (2004). *Cognitive neuroscience of attention*. New York: Guilford Press.

Rasmussen, C., & Bisanz, J. (2005). Representation and working memory in early arithmetic. *Journal of Experimental Child Psychology, 91,* 137- 157.

Reuhkala, M. (2001). Mathematical Skills in Ninth-graders: Relationship with visuospatial abilities and working memory. *Educational Psychology, 21,* 387-399.

Roodenrys, S., Hulme, C., & Brown, G. (1993). The development of short-term memory span: Separable effects of speech rate and long-term memory. *Journal of Experimental Child Psychology, 56*, 431-442.

Ross-Sheehy, S., Oakes, L.M., & Luck, S.J. (2003). The development of visual short-term memory capacity in infants. *Child Development, 74,* 1807-1822.

Ruckridge, J.R., & Tannock, R. (2002). Neuorpsychological profiles of adolescents with ADHD; effects of reading difficulties and gender. *Journal of Child Psychology and Psychiatry, 43,* 988-1003.

Saults, J.S., & Cowan, N. (1996). The development of memory for ignored speech. *Journal of Experimental Child Psychology, 63*, 239-261.

Schneider, W., Kron, V., Hunnerkopf, M., & Krajewski, K. (2004). The development of young children's memory strategies: First findings from the Würzburg Longitudinal Memory Study. *Journal of Experimental Child Psychology, 88*, 193-209.

Siegel, L.S. (1988). Evidence that IQ scores are irrelevant to the definition and analysis of reading-disability. *Canadian Journal of Psychology, 42,* 201-215.

Siegel, L.S., & Ryan, E.B. (1989). The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development, 60,* 973-980.

Siegler, R. S. (1996). *Emerging minds: The process of change in children s thinking*. New York: Oxford University Press.
Simon, H.A. (1974). How big is a chunk? *Science, 183*, 482-488.
Smith, L.B., Thelen, E., Titzer, R., & McLin, D. (1999). Knowing in the context of acting: The task dynamics of the A-not-B error. *Psychological Review, 106*, 235-260.
Sokolov, E.N. (1963). *Perception and the conditioned reflex*. NY: Pergamon Press.
Sonuga-Barker, E.J.S., Dalen, L., Daley, D., & Remington, B. (2002). Are planning, working memory, and inhibition associated with individual differences in preschool ADHD symptoms? *Developmental Neuropsychology, 21*, 255-272.
Stothard, S. E., & Hulme, C. (1992). Reading comprehension difficulties in children. *Reading and Writing: An Interdisciplinary Journal, 4,* 245-256.
Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662.
Swanson, H.L. (2003). Age-related differences in learning disabled and skilled readers' working memory. *Journal of Experimental Child Psychology, 85*, 1-31.
Swanson, H.L. & Beebe-Frankenberger, M. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for math disabilities. *Journal of Education Psychology, 96*, 471–491.
Swanson, H.L., & Sachse-Lee, C. (2001). Mathematical problem solving and working memory in children with learning disabilities: Both executive and phonological processes ar e important. *Journal of Experimental Child Psychology, 79*, 294-321.
Swanson, H.L. & Saez, L. (2003). Memory difficulties in children and adults with learning disabilities. In H.L. Swanson, *S. Graham*, & *K.R. Harris* (Eds.), *Handbook of learning disabilities*, pp. 182-198. New York: Guildford Press.
Thevenot, C., & Oakhill, J. (2005). The strategic use of alternative representations in arithmetic word problem solving. *Quarterly Journal of Experimental Psychology, 58A,* 1311-1323.
Todd, J.J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature, 428,* 751-754.
Towse, J.N., Hitch, G.J., & Hutton, U. (1998). A reevaluation of working memory capacity in children. *Journal of Memory and Language*, 39, 195-217.
Treisman, A. (1988). Features and objects: The fourteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology, 40A*, 201-237.
Tulving, E., & Patkau, J.E. (1962). Concurrent effects of contextual constraint and word frequency on immediate recall and learning of verbal material. *Canadian Journal of Psychology, 16*, 83-95.
Turner, M.L., & Engle, R.W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28*, 127-154.
Unsworth, N., & Engle, R.W. (In press). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*.
Vogel, E.K., McCollough, A.W., & Machizawa, M.G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature, 438,* 500-503.
Wagner, R. K, & Muse. A. (in press). Working memory deficits in developmental dyslexia. In T. P. Alloway, T. P. & S. E. Gathercole (Eds), *Working memory in neurodevelopmental conditions*. Psychology Press.

Wagner, R., Torgesen, J., & Rashotte, C. (1999). *Comprehensive Test of Phonological Processing.* Austin, TX: Pro-Ed.

Wagner, R. K., Torgesen, J.K., & Rashotte, C.A., Hecht, S.A., Barker, T.A., Burgess, S.R., Donahue, J., & Garon, T. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental Psychology, 33*, 468-479.

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

Weiss, V. (1995). Memory span as the quantum of action of thought. *Cahiers de Psychologie Cognitive, 14*, 387-408.

Xu, Y., & Chun, M.M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, *440*, 91-95.

Figure Captions

*need permission

*Figure 1. Three conceptual, behavioral models of working memory. From Cowan (2005), *Working memory capacity*. New York: Psychology Press (p. 18, Figure 1.2)

*Figure 2. The relation between speech rate and memory span in 4-year-old and 8-year-old children. From 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 and Language, 33*, 234-250. (p. 240, Figure 1)

*Figure 3. Span and post-span (ability-adjusted) trials within three phases of an experiment in which one of two groups of children received instructions to speak quickly in Phase 2. Top panels show speech rates and bottom panels show memory scores. From Cowan et al. 2006, Psych Science speed and span (simplify) Cowan, N., Elliott, E.M., Saults, J.S., Nugent, L.D., Bomb, P., & Hismjatullina, A. (2006). Rethinking speed theories of cognitive development: Increasing the rate of recall without affecting accuracy. *Psychological Science*, *17*, 67-73. (p. 70, Figure 2)

*Figure 4. In three age groups, decay of memory in the final serial position of lists unattended during their presentation. The lists were adjusted based on span for attended lists and no decay differences were found elsewhere in the lists. From Cowan, N., Nugent, L.D., Elliott, E.M., & Saults, J.S. (2000). Persistence of memory for ignored lists of digits: Areas of developmental constancy and change. *Journal of Experimental Child Psychology, 76*, 151-172. (p. 164, Figure 3, top panel)

*Figure 5. Memory for speech that was attended (dashed lines) or ignored (solid lines) during its presentation, for three age groups. From Cowan, N., Elliott, E.M., Saults, J.S., Morey, C.C., Mattox, S., Hismjatullina, A., & Conway, A.R.A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*, 42-100. (p. 55, Figure 1)

*Figure 6. Modified versions of three complex span tasks: the counting span (Case et al., 1982), sentence span (Daneman & Carpenter, 1980), and operation span (Turner & Engle, 1989) tasks. Note that Engle and colleagues have promoted a version in which the item to be recalled is a word presented after each processing episode that did not come from it.

*Figure 7. Mean response durations for three working memory tasks in three age groups. From Cowan, N., Towse, J.N., Hamilton, Z., Saults, J.S., Elliott, E.M., Lacey, J.F., Moreno, M.V., & Hitch, G.J. (2003). Children's working-memory processes: A response-timing analysis. *Journal of Experimental Psychology: General*, *132*, 113-132. (p. 124, Figure 3)

*Figure 8. The relation between the rate of speech and memory span for words and nonwords. From Hulme, C., Maughan, S., & Brown, G.D.A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory & Language*, *30*, 685-701. (p. 690, Figure 1)

Figure 9. An illustration of the two presentation conditions of Cowan, Saults, & Morey (2006). Each number represents the serial order of a name presented within the location shown by a stylized house. There could be 3 to 7 items per trial and there always was the same number of names and locations. In the uneven condition (bottom panel), at least one location was used twice, decreasing the number of locations used but adding to the spatial path recursion, and therefore confusion. The 1-to-1 condition produced better performance in adults with no dual task; the uneven condition produced better performance in third-grade children, and also in adults carrying out articulatory suppression during the task.

*Figure 10. Pattern of performance on 1-to-1 trials and on uneven trials for which 1 versus 2 names were assigned to the house that serves as the correct answer. Adults with articulatory suppression perform almost exactly like third-grade children. From Cowan, N., Saults, J.S., & Morey, C.C. (2006). Development of working memory for verbal-spatial associations. *Journal of Memory and Language*, *55*, 274-289. (p. 282, Figure 3).

Figure 1

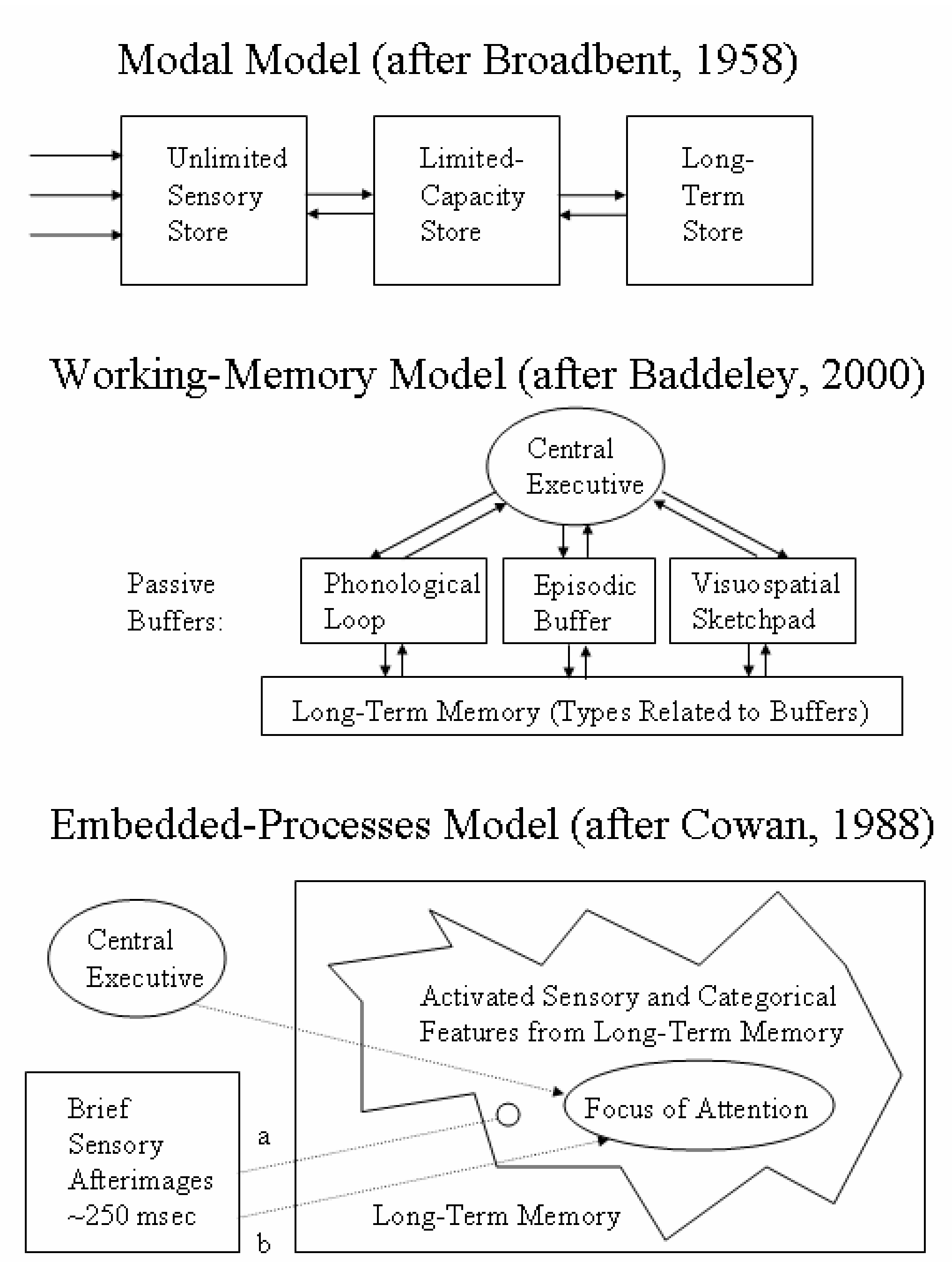

Modal Model (after Broadbent, 1958)
Unlimited Sensory Store
Limited-Capacity Store
Long-Term Store
Working-Memory Model (after Baddeley, 2000)
Central Executive
Passive Buffers:
Phonological Loop
Episodic Buffer
Visuospatial Sketchpad
Long-Term Memory (Types Related to Buffers)
Embedded-Processes Model (after Cowan, 1988)
Central Executive
Activated Sensory and Categorical Features from Long-Term Memory
Focus of Attention
Brief Sensory Afterimages ~250 msec
a
b
Long-Term Memory

Figure 2

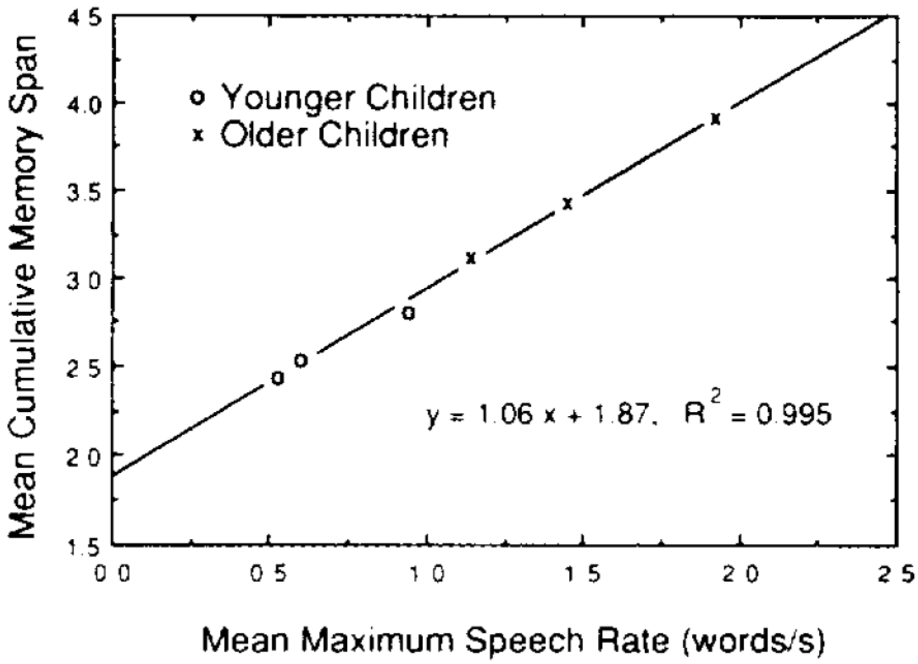

Mean Cumulative Memory Span
o Younger Children
x Older Children
y = 1.06 x + 1.87, R² = 0.995
4.5
4.0
3.5
3.0
2.5
2.0
1.5
0.0
0.5
1.0
1.5
2.0
2.5
Mean Maximum Speech Rate (words/s)

Figure 3

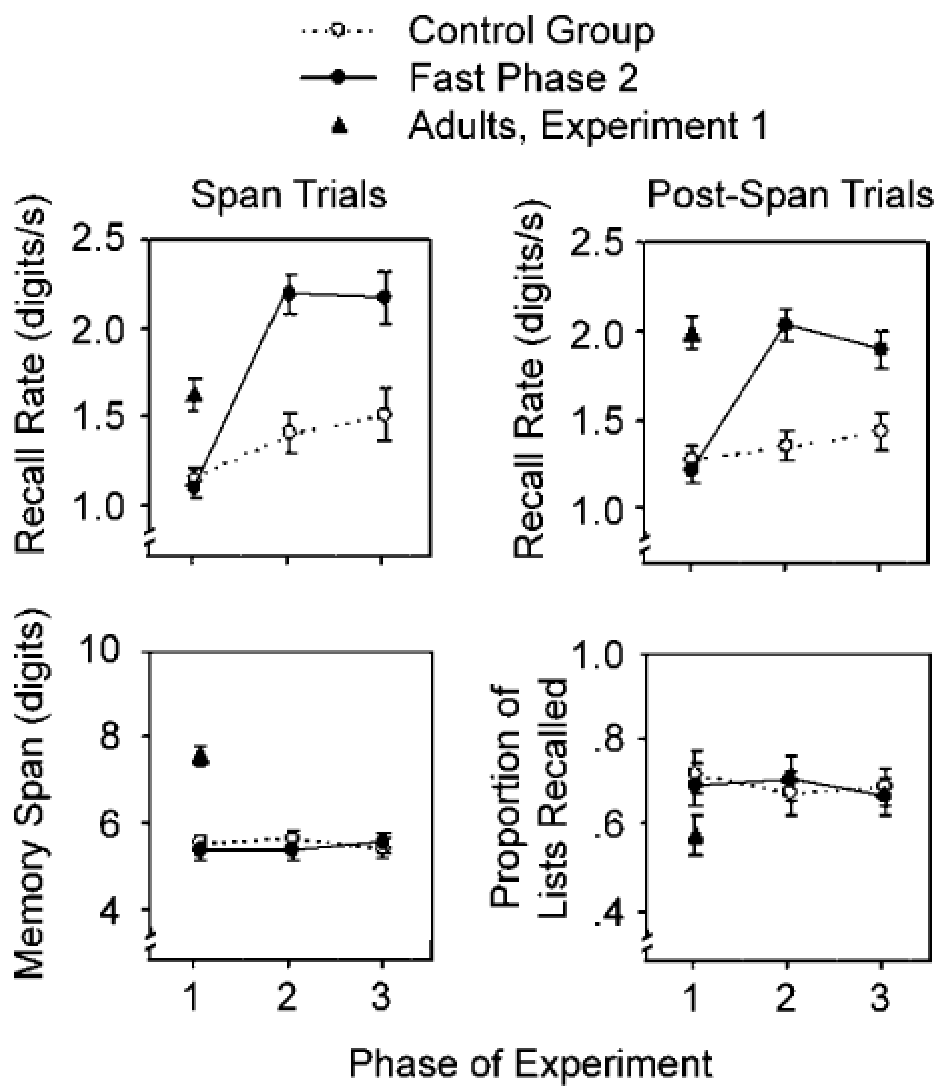

Control Group
Fast Phase 2
Adults, Experiment 1
Span Trials
Post-Span Trials
Recall Rate (digits/s)
2.5
2.0
1.5
1.0
Memory Span (digits)
10
8
6
4
Proportion of Lists Recalled
1.0
.8
.6
.4
1
2
3
Phase of Experiment

Figure 4

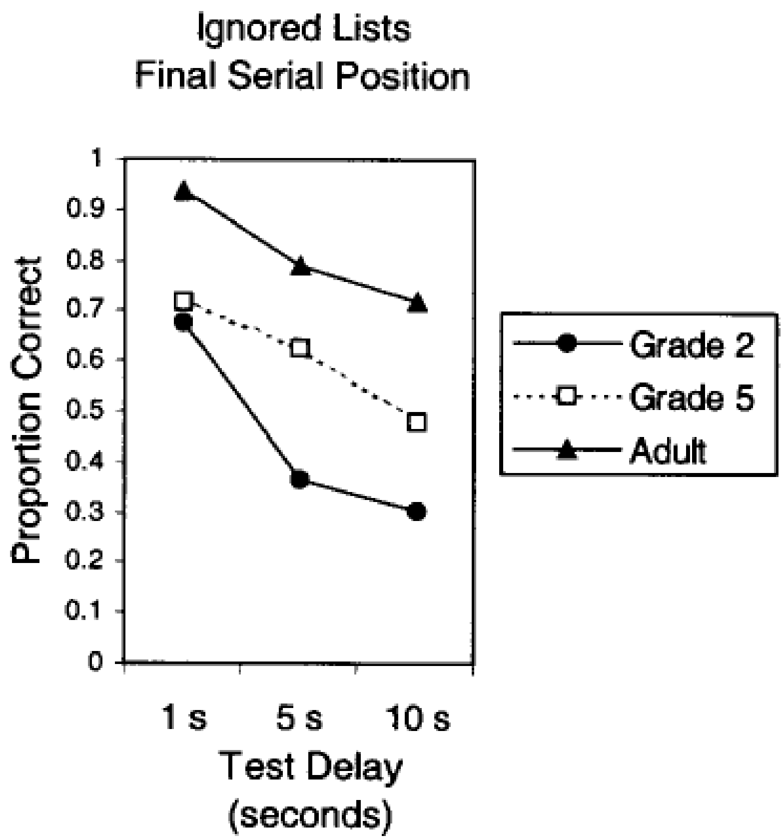
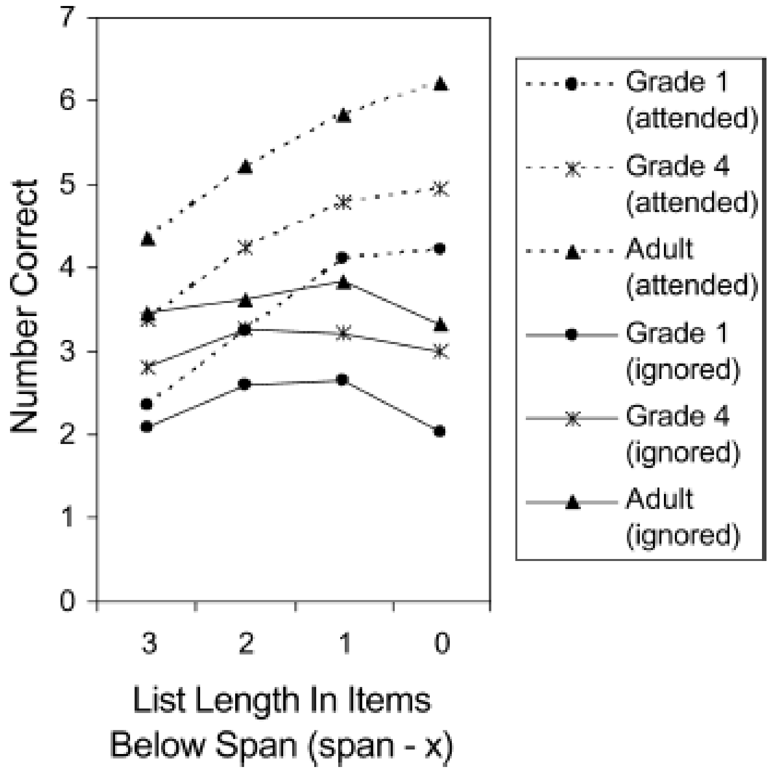

Ignored Lists
Final Serial Position
Proportion Correct
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1 s
5 s
10 s
Test Delay
(seconds)
Grade 2
Grade 5
Adult

Figure 5



7
6
5
4
3
2
1
0
Number Correct
3
2
1
0
List Length In Items
Below Span (span - x)
Grade 1 (attended)
Grade 4 (attended)
Adult (attended)
Grade 1 (ignored)
Grade 4 (ignored)
Adult (ignored)

Figure 6

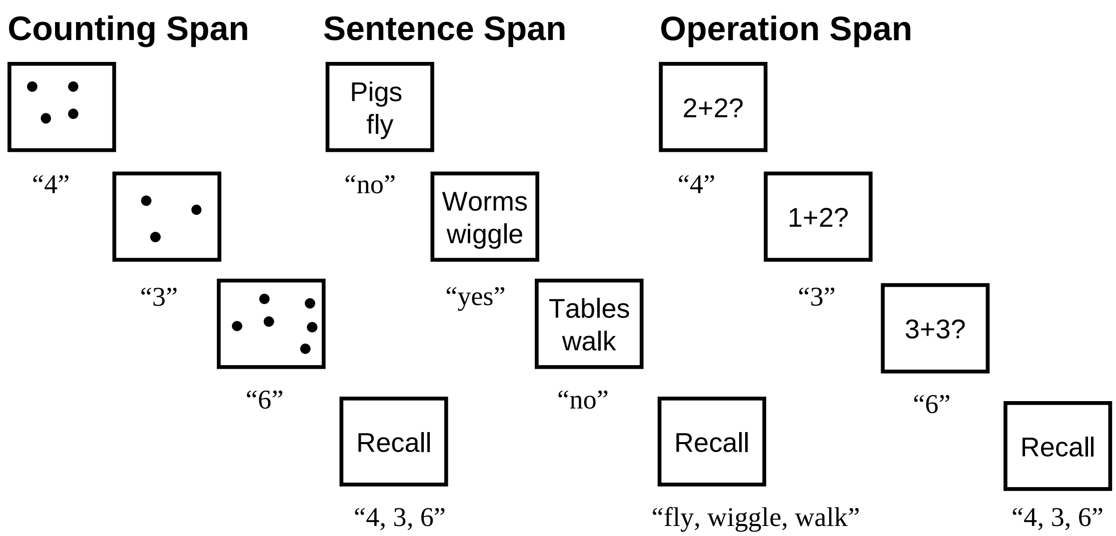

Counting Span
Sentence Span
Operation Span
"4"
"3"
"6"
Recall
"4, 3, 6"
Pigs fly
"no"
Worms wiggle
"yes"
Tables walk
"no"
Recall
"fly, wiggle, walk"
2+2?
"4"
1+2?
"3"
3+3?
"6"
Recall
"4, 3, 6"

Figure 7

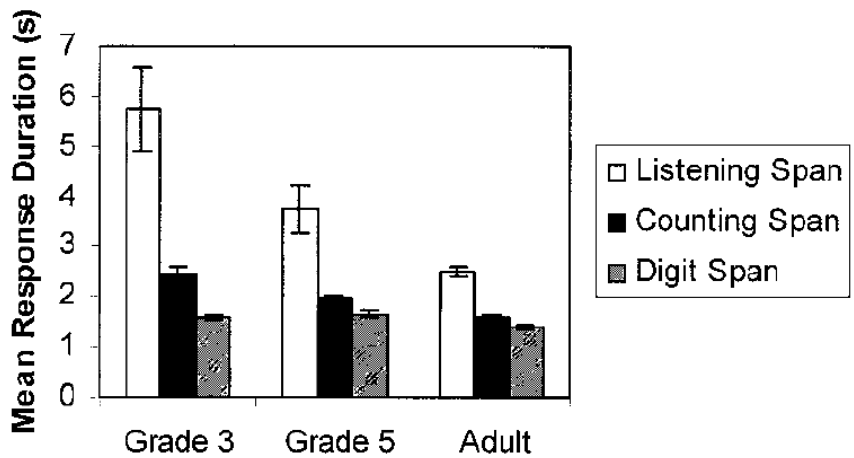

Mean Response Duration (s)
0
1
2
3
4
5
6
7
Grade 3
Grade 5
Adult
Listening Span
Counting Span
Digit Span

Figure 8

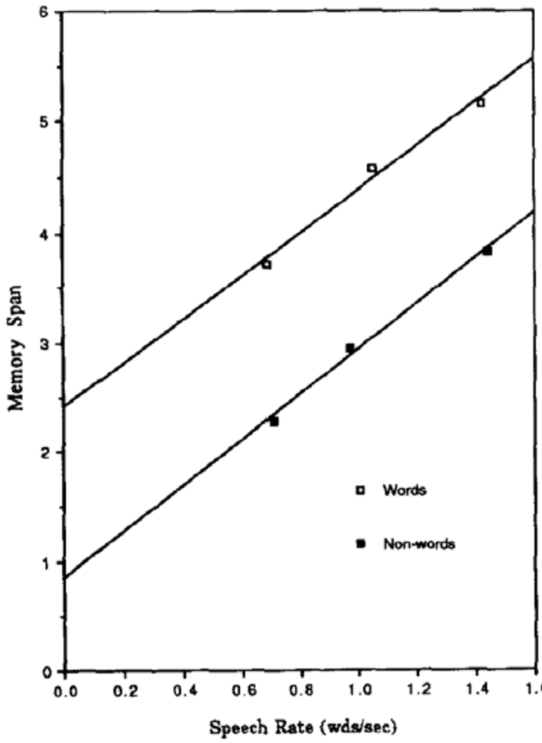
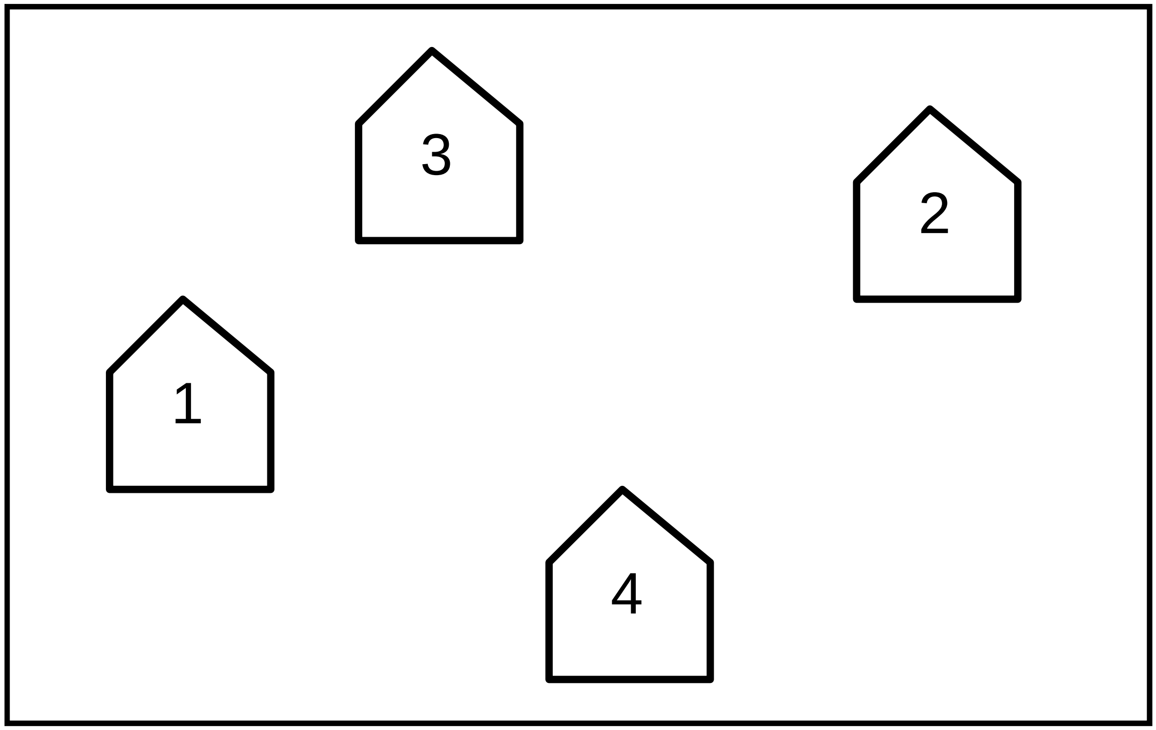

Memory Span
Speech Rate (wds/sec)
Words
Non-words
0
1
2
3
4
5
6
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6

Figure 9

1-to-1 Condition



Uneven Condition

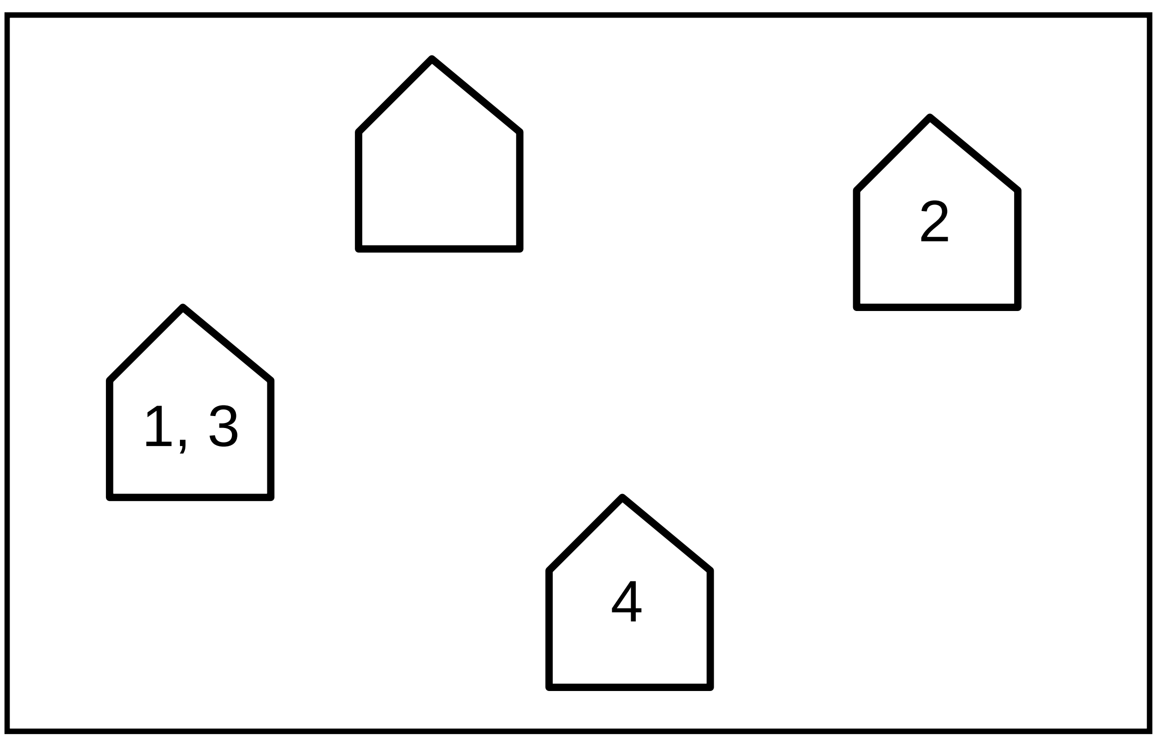
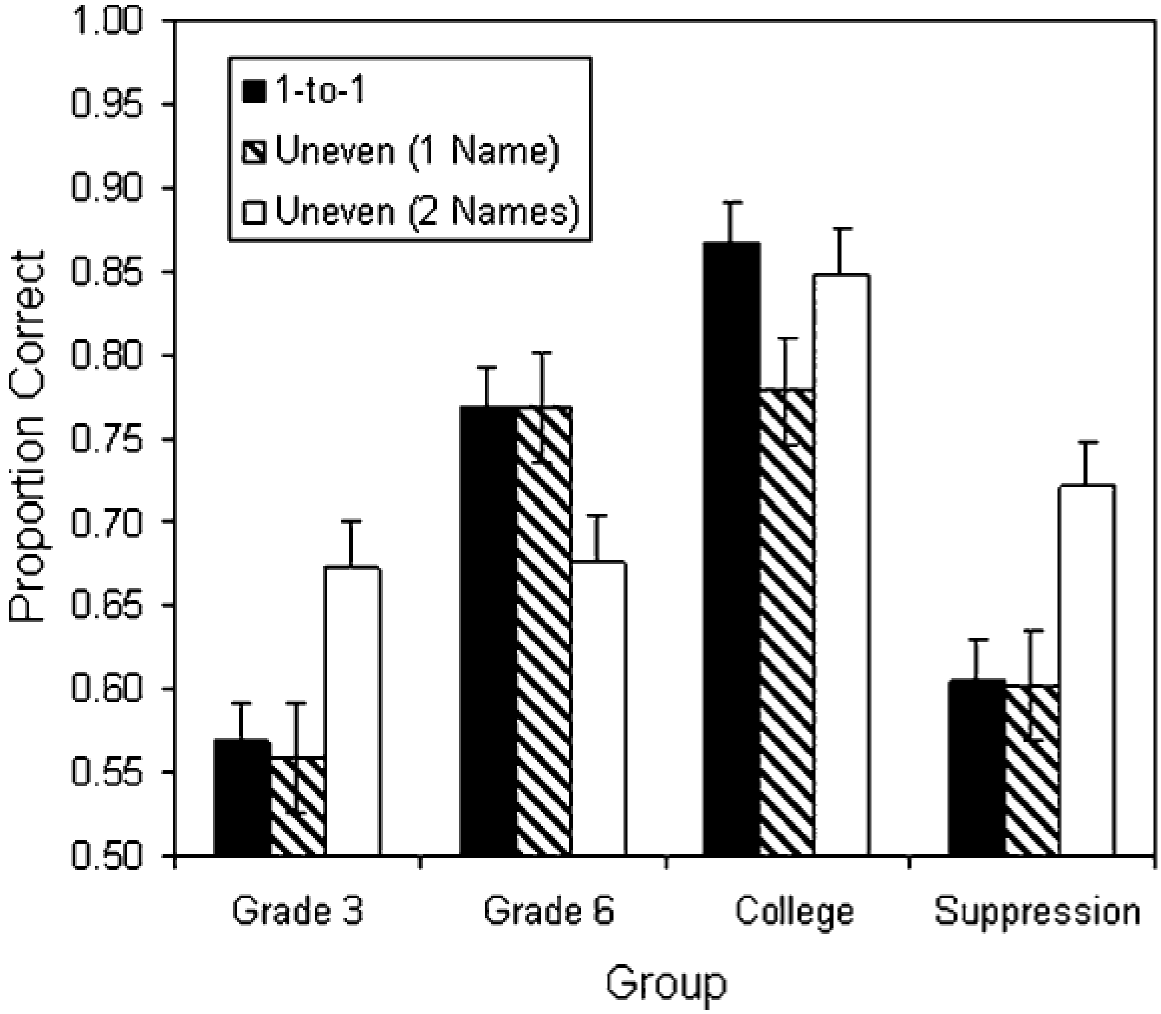


Figure 10



1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
Proportion Correct
1-to-1
Uneven (1 Name)
Uneven (2 Names)
Grade 3
Grade 6
College
Suppression
Group