

2002

Advances in Child
Development and
Behavior

Vol. 29

THE SEARCH FOR WHAT IS FUNDAMENTAL IN THE DEVELOPMENT OF WORKING MEMORY

Nelson Cowan,¹ J. Scott Saults, and Emily M. Elliott

DEPARTMENT OF PSYCHOLOGICAL SCIENCES
UNIVERSITY OF MISSOURI
COLUMBIA, MISSOURI 65211

- I. INTRODUCTION
 - A. DEFINITION AND DESCRIPTION OF WORKING MEMORY
 - B. EMPHASIS ON ELEMENTARY OPERATIONS
 - C. A MODEL OF PROCESSING IN WORKING MEMORY TASKS
 - D. THE STATUS OF THE FIELD OF WORKING MEMORY DEVELOPMENT IN RELATION TO THE PRESENT APPROACH
 - E. DIFFICULTY-INSENSITIVE MEASURES OF DEVELOPMENTAL CHANGE
- II. EVIDENCE OF DEVELOPMENTAL CHANGE IN PROCESSING PARAMETERS OF WORKING MEMORY
 - A. CAPACITY LIMITS
 - B. RETRIEVAL RATES
 - C. DECAY RATES
- III. CONCLUSIONS, OBSERVATIONS, AND SPECULATIONS
 - A. ARE THERE TRULY DIFFICULTY-INSENSITIVE MEASURES OF WORKING-MEMORY PARAMETERS?
 - B. WHAT ARE THE RELATIONS BETWEEN THE STUDIED WORKING-MEMORY PARAMETERS?
 - C. WHAT ARE THE BASES OF DEVELOPMENTAL CHANGE IN THESE WORKING-MEMORY PARAMETERS?
 - D. HOW SUCCESSFUL ARE THESE WORKING-MEMORY PARAMETERS IN ACCOUNTING FOR HIGHER LEVEL COGNITION?
 - E. WHAT ARE THE IMPLICATIONS FOR OTHER MODELS OF WORKING MEMORY AND ITS DEVELOPMENT?

REFERENCES

¹Address correspondence to Nelson Cowan, Department of Psychological Sciences, University of Missouri, 213 McAlester Hall, Columbia, MO 65211, USA. E-mail: CowanN@missouri.edu.

I. Introduction

A. DEFINITION AND DESCRIPTION OF WORKING MEMORY

Given the vast storehouse of one's knowledge, a mechanism is needed to hold a limited amount of especially relevant information in a privileged state for a short time. That mechanism, or collection of mechanisms (Cowan, 1999a), is generally implied when investigators speak of working memory. Although definitions of working memory vary dramatically from one investigator to another (Miyake & Shah, 1996), they all share a core idea (e.g., Baddeley, 1986). To carry out just about any cognitive task, a limited amount of information must be kept available temporarily. For example, comprehension logically requires some access to the information from early in a sentence, in some form, to be held until it can be integrated with information from later in the sentence, resulting in a coherent linguistic and semantic interpretation. Similarly, problem-solving logically requires access to data, premises, and so on until they can be integrated to produce a solution. Developmental differences in working memory, thus defined, certainly must play an important role in the growth of intellectual and scholastic abilities. In the present chapter, the development of working memory between the age of 7 years and adulthood is examined, largely on the basis of some of our own cognitive developmental studies.

The information in working memory, presumably, can come from the outside world but it also can include information from long-term memory that becomes activated as a result of the task context and the individual's attempt to carry out the task (Cowan, 1993, 1995, 1999a). Part of the difficulty in defining working memory is that it is not confined to a specific set of tasks; almost any cognitive task can involve the temporary activation of information that is needed for the response (Shiffrin, 1993). The question then becomes how one can logically identify and characterize the basic mechanisms that are used to activate information, and hold it temporarily, across such a wide range of cognitive situations. That question is compounded by the difficulty of understanding the development of working memory, given the need to compare basic mechanisms fairly across children of different ages, who are at different performance levels.

B. EMPHASIS ON ELEMENTARY OPERATIONS

When the title of this chapter was first formulated, we drew inspiration from an earlier work by Posner and McLeod (1982) entitled "Information Processing Models—In Search of Elementary Operations." That work nicely set the stage for our analytic investigation as follows (Posner & McLeod, 1982, p. 478):

In a criticism of cognitive psychology, Newell (1973) suggested that there are few common principles emerging from myriads of often dichotomous distinctions. Newell's remedy . . . is the construction of complex information processing models that might eventually simulate a wide range of human mental activity. The reverse emphasis, and the one explored in this chapter, is that detailed studies of particular task configurations will lead to the identification of fundamental operations that can be used to characterize the human mind.

A search for such elementary operations is, we believe, the most important initial route for an understanding of working memory. However, we also think that we have learned enough in the past few years to hazard a guess at a type of information processing model, as well. Because such a model can make it easier to think about empirical results, we first describe the model, under I.C. Then we introduce, for comparison, the status of the field of working memory development, under I.D., and an aspect of developmental methodology and metatheory that we emphasize, under I.E. After that, under Part II, we pursue this analytic approach to indicate how parts of our model have been examined and which ones remain to be confirmed.

C. A MODEL OF PROCESSING IN WORKING MEMORY TASKS

The present theoretical framework is based on a range of supporting evidence (see Cowan, 1988, 1993, 1995, 2001a) beyond the scope of the present chapter. In this framework, nodes from the memory system are automatically activated by incoming stimuli, although this automatically activated portion of memory consists mostly of physical features, i.e., sensory memory. This sensory memory decays (and is degraded through interference from other stimuli) within a matter of seconds unless it becomes the focus of the participant's attention. Once in the focus of attention, additional processing of the sensory information occurs and more categorical features related to the stimulus become activated. (Once they leave the focus of attention, these categorical features also suffer from interference or decay.) The focus of attention is limited in capacity and the rate at which activated nodes can be retrieved into the focus of attention also is limited. Presumably, these limits are important because deliberate recall can occur only on the basis of the current contents of the focus of attention.

Although this description is incomplete, it highlights the importance of three basic processing parameters for which we have invested considerable effort in order to find valid measures that could be used in developmental studies. These three parameters are the capacity of the focus of attention, the rate of retrieval of activated information into this focus, and the rate of decay of information from sensory memory. These are illustrated in Fig. 1.

The brief sketch of processing shown in Fig. 1 leaves a number of finer points to be discussed.

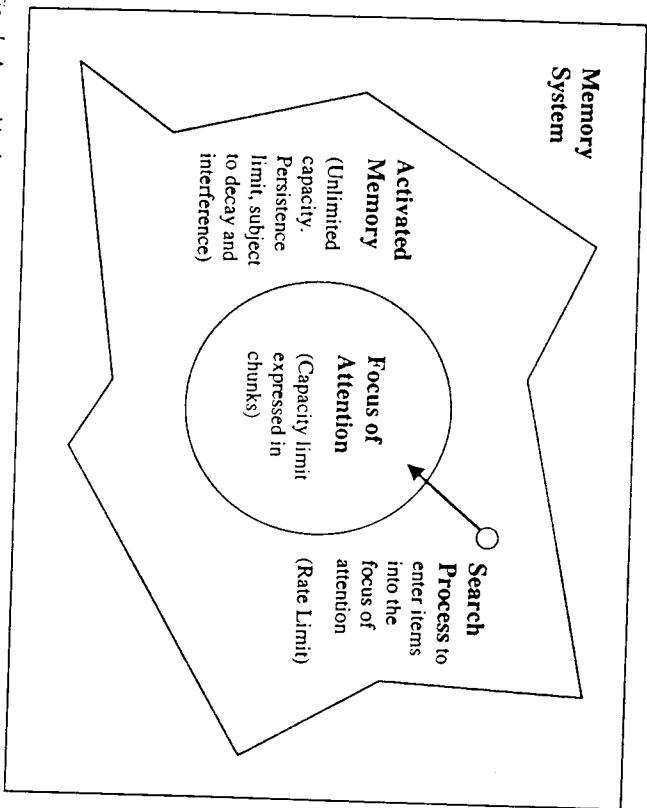


Fig. 1. A graphic depiction of various components of the memory system and their relation to the working memory limits we discuss. Activated memory and the focus of attention together form the working memory system according to this view.

1. Feature-Binding Processes

One particularly important point is that working memory must contain more than the activated or attended information from long-term memory. If that were all, no new information could be learned or recalled. Working memory must keep track of which elements are activated at the same time, which ones were activated after which other ones, and so on. In other words, the binding between activated features must be saved. Although Cowan (1988, 1995, 1999a) has often been cited to the effect that working memory is just the activated portion of long-term memory, the later two references actually cautioned otherwise. Cowan (1995, p. 101; 1999, p. 89) proposed that new links between concurrently activated elements also can be formed and stored in memory, with the assistance of attention. These new links are involved in the preservation of serial order and serial position information and other episodic information, as well as in the combination of elements or features to form novel objects. Such a role of attention is compatible with the role assigned to it by the feature conjunction model (e.g., Treisman, 1987).

2. New Long-Term Storage

Additionally, we suggest that, after attended sets of features are grouped into objects (i.e., chunks), only about four of these objects can be held in the focus of

attention at once (with the number subject to individual differences and age differences). However, once associations between items are formed (whether strongly enough to create discrete, multi-item chunks, or more weakly to form some more complex structure), the associative structure can be saved in long-term memory. Later, these new structures can be reactivated in a manner similar to preexisting information in memory.

3. Focal-Attention Zoom Lens

In a theoretical review, Cowan (2001a) proposed that normal adults' average capacity of the focus of attention is about four unconnected chunks. However, several commentaries on that review suggested that sometimes the focus of attention can cover as little as one chunk. McElree and Doshier (2001) described research from a probed reaction time procedure in which response accuracy was examined for cases in which a list was presented, followed by a probe that matched any one item in the list or did not match any items, a procedure like that of Sternberg (1966). In their revised procedure, however, a variable response deadline was given. Such deadlines can reveal the speed of memory retrieval, or "retrieval dynamic," that is the function of accuracy increasing as the response deadline grows more lax. The results showed a faster retrieval dynamic for the most recent list item than for previous items, which did not differ from one another. This pattern could be explained on the basis that only the last item actually was in the focus of attention and therefore could be recalled without first having to be retrieved from some other portion of the memory system outside of attention. Moreover, Usher, Haarmann, Cohen, and Horn (2001) described a cognitive model in which the focus of attention could zoom from one to four depending on the task circumstances. In the response to commentaries, Cowan (2001b) tentatively accepted this view that the focus of attention can vary between a more intensive mode and a wider mode.

As in Usher's model, the limit of breadth is about four chunks. In Usher's model, this capacity limit occurs because the representations of different chunks overlap. Without such overlap, important cognitive functions such as generalization would be impossible. With such overlap, the model is unable to represent more than four chunks simultaneously. Cowan (2001a) and some of the other commentaries discussed a related neural representation in which each item in the focus of attention is represented by simultaneously pulsing feature detectors for the various features in the object. Multiple objects must pulse at different moments to avoid confusion between them but all objects in attention must pulse at a certain minimal rate to stay active. With too many objects in attention at once, the objects' features become temporally confused with one another. This places an upper, but not a lower, limit on how much can be in the focus of attention at once. Presumably, the fewer the number of simultaneously held stimulus objects (or chunks), the less the chance of confusion between them. Also, with fewer chunks from the stimulus set in the

focus of attention at one time, it should be easier to carry out intensive processing by incorporating other chunks from long-term memory.

4. Processing in Serial Recall

We now show how these theoretical claims about processing apply to one relatively simple task, the serial recall task, given our reliance on this task in the present chapter.

a. Attention and Formation of New Associative Structure. New associations can be formed quickly to assist performance even in short-term recall. In a simple task in which a list of items (e.g., digits) is to be recalled, one need not assume that the participant holds all of the items in the focus of attention at once. Instead, a newly formed, active structure may be held automatically as the participant focuses attention on only a portion of the structure at a time to recall it.

b. Attention and Phases of Processing. Such a process solves another problem, which is the following: If attention leads to the grouping of items into chunks and adults can hold up to about four chunks at once in the focus of attention, then why does the response protocol in short-term serial recall include delays, both in children (e.g., Cowan, 1992) and in adults (Hulme et al., 1999; Tehan & Lalor, 2000)? Moreover, if the chunks can be independent, how does the participant know which one to recall first, second, and so on? The zoom feature of attentional capacity, described above, can help to explain all of this. Perhaps one must focus attention entirely on a chunk in order to recall it overtly (e.g., to pronounce a word from a stimulus list), rather than spreading attention across all of the chunks simultaneously.

Thus, the list recall process would have several phases, as illustrated in Fig. 2. In the first, during input of the list, the scope of attention would vacillate. It could zoom in for the complete encoding of a presented item (*item encoding*) and zoom out again to encompass several recently activated items in order to form new associative links between them or between items and serial positions (*item binding*). In the next, *preparatory* phase, the attentional focus would spread out in order to get an overview of the entire, newly formed associative structure of the list. In the *search and output* phase, attention would typically be narrowed to a single chunk (group of items) every time it was about to be recalled. During the interword pauses between words in a response, a narrow focus of attention could scan the associative structure to determine which item to recall next; or, alternatively, the focus could zoom out to look at the list and find out which one to recall next and then zoom in to execute recall of the selected item. In either case, an assumption of a correlation between list length and the amount of associative structure or material to be searched correctly predicts the finding that longer lists result in longer interword pauses (Cowan, 1992; Cowan et al., 1994, 1998, 1999; Hulme et al., 1999; Tehan & Lalor, 2000).

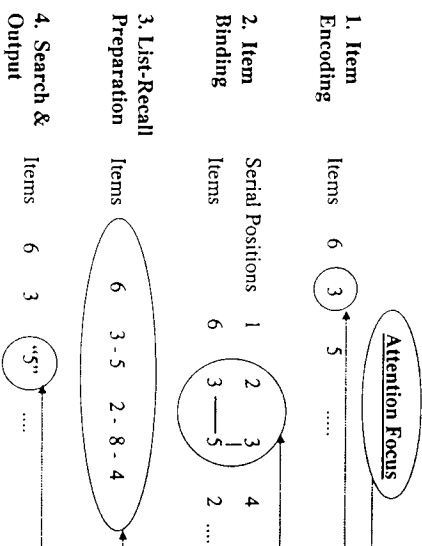


Fig. 2. An illustration of how the scope of the focus of attention (depicted as a circle or ellipse) might change from one phase of a memory trial to another. The depicted trial involves the serial recall of digits. Attention may focus on a single item during its encoding and then zoom out to allow an associative binding between nearby items and between an item and its serial-position context and in again to encode the next item and so on. At the end of the list, it may zoom out further to reload the series of chunked items into the focus of attention in preparation for recall, which may allow the creation of a more organized, overall mental scheme for the list before it is recalled. The process of recall may require that the focus of attention zoom in again, a requirement that could account for the well-known phenomenon of output interference.

c. Involvement of Retrieval and Decay Rates. Whenever an item could not be rehearsed, for example during the overt recall of another item, the unrehearsed item would tend to be forgotten over time. We broadly refer to that forgetting as decay, without specifying its exact nature. The amount of decay may be related to the absolute amount of time that has passed without rehearsal or, alternatively, to the amount of time relative to the interitem intervals, a loss of distinctiveness of the most recent items with the passage of time (Cowan, Saults, & Nugent, 1997, 2001; Nairne, Neath, Serra, & Byun, 1997). In any case, faster retrieval of information (because of faster control of the focus of attention or other, unknown factors) limits the amount of time during which decay can take place. Thus, in a serial recall task, the capacity of attention, the short-term memory search or retrieval rate, and the decay rate all could come into play in determining performance. Cowan (2001a,b) discussed the application of the framework to other tasks.

D. THE STATUS OF THE FIELD OF WORKING MEMORY DEVELOPMENT IN RELATION TO THE PRESENT APPROACH

Before summarizing research on development of working memory parameters, we want to convey briefly what progress has been made in other research on the development of working memory. The present approach is similar to conventional

approaches in some ways and different in other ways, and these similarities and differences must be described and put in the correct context.

1. *Working Memory as a Distinct Research Area*

Some theorists consider memory to be just one type of problem for which the child must find a cognitive solution. Winness, for example, an article in which the possible demise of memory as a research area was proclaimed (Kuhn, 2000). This general approach to memory would not suffice to explain working memory. If one is dealing with the boundless well of information that the individual has learned over a lifetime, then the structure of information in the memory representation may be of similar importance for any type of cognitive task, mnemonic or otherwise. However, for any of these tasks, assuming that a limited amount of information becomes temporarily accessible (as the concept of working memory seems to imply), one can discuss certain fundamental information-processing limits: how quickly a limited amount of information moves into a temporarily accessible state, how much information can become accessible in this way concurrently, and how long it stays accessible. If information accessibility is limited in this manner, as we believe, these limits are important to understand apart from the knowledge structure for a particular task.

2. *Approaches to Working Memory Development*

a. Modular Approach. Another key point concerns the procedures for examining working memory and their application to developmental research. The approach initiated by Baddeley and Hitch (1974) and elaborated upon enormously since that time (e.g., Baddeley, 1986) delineates certain hypothetical structures of working memory referred to as a central executive, a phonological loop, and a visuospatial sketchpad. The emphasis has been on learning how these components work, with a subsidiary emphasis on learning how the components mediate important types of learning (Baddeley, Gathercole, & Papagno, 1998; Gathercole & Baddeley, 1993). Some investigators have applied the same type of theoretical analysis to an understanding of developmental change (e.g., Henry, 1991; Hitch, Halliday, Dodd, & Litter, 1989; Hulme & Tordoff, 1989).

b. Psychometric Approach. In contrast to the modular approach initiated by Alan Baddeley and his colleagues, another approach has used more complex, holistic working memory measures that capture more of the variance of complex scholastic abilities tests (Daneman & Carpenter, 1980; Daneman & Merikle, 1996), presumably because they require both the processing of information and the storage of information during that processing. Researchers have attempted to analyze these complex working memory tasks into their basic components, both by examining the performance of high- and low-span individuals on various other, more narrowly

defined cognitive tasks (Engle, Conway, Tuholski, & Shisler, 1995; Conway & Engle, 1994) and by using structural equation models to isolate the source of variance within the complex working memory tasks that accounts for general intelligence (Engle, Tuholski, Laughlin, & Conway, 1999b). Some investigators have applied that type of task to an understanding of developmental change also (e.g., Hitch, Towse, & Hutton, 2001; Towse, Hitch, & Hutton, 1998).

Both of these approaches in the field (modular and psychometric) combine analytic experimental work and correlational work, though the balance of these two research emphases differs for the two types of approach, with more correlational work in the psychometric approach. Within the psychometric approach, the debate concerns whether individual and age differences in working memory stem from differences in the use of attention and inhibition or from rates of retrieval and forgetting. In the first camp, one would place theoretical views holding that the critical variable is the control of attention (Engle *et al.*, 1999b) or the use inhibition (Hasher, Stolzhus, Zacks, & Rypina, 1991). In the second camp, one would place views holding that the critical variable is the speed of processing (Kail, 1992; Kail & Salthouse, 1994; Salthouse, 1996) or the additional forgetting that can occur over time as a result of slower processing and responding (Hitch *et al.*, 2001; Towse *et al.*, 1998). The former view has been taken to imply that attention is shared between processing and storage and that working-memory limits stem from difficulty doing this sharing, whereas the latter view has been taken to imply that attention switches back and forth between storage and processing (see Hitch *et al.*, 2001).

3. *The Present Approach Compared to the Others*

a. Comparison to the Modular Approach. The present approach illustrated in Fig. 1 borrows from both the modular and psychometric approaches, but adds something new. From the modular approach, it borrows the assumption that working memory results from an interaction between active attentional processes and passive memory-storage mechanisms and borrows the notion that we need to isolate, identify, analyze, and characterize these mechanisms. In the present approach, no distinction is drawn between phonological and visuospatial subsystems because we still do not know the best taxonomy. We do not know, for example, whether memory for nonverbal sounds or tactile sensations can be considered part of either of these stores. We do not know if memory for, say, spatial arrays of words must engage two separate, phonological and spatial subsystems or one integrated system. In the present approach, we circumvent such questions by simply referring to all such memory, whether physical or categorical in nature, as memory activation. Similar dynamic principles may apply to various types of activation, for example similar parameters of decay and of interference, modulated by the extent to which subsequent stimuli share similar features (Cowan & Saults, 1995; Nairne, 1990).

Baddeley (1986) spoke of passive memory stores as if they did not undergo developmental change in the forgetting rate. This was a parsimonious account because it allowed him to attribute developmental change entirely to the efficiency with which mnemonic processing can be carried out before the passive memory is lost (cf. Case, Kurland, & Goldberg, 1982). However, we have shown subsequently that passive memory at least for acoustic information is lost more quickly in younger children (Cowan *et al.*, 2000; Keller & Cowan, 1994; Gomes *et al.*, 1999; Saults & Cowan, 1996).

Unlike the approach of Baddeley (1986), we also consider that the attentional process itself has storage capability. Baddeley (2000) has modified his account to include an episodic short-term buffer that depends on attentional processes for its input; so he may not differ from us much on this score. He has conjectured that his episodic buffer has the capacity limit that we ascribe to the focus of attention (Baddeley, 2001). He has suggested that his buffer cannot consist entirely of the activated portion of long-term memory as it also must include newly formed links between active concepts and, although he has contrasted that with the Cowan model, Cowan (1995, 1999a, 2001b) actually made a similar point.

b. Comparison to the Psychometric Approach. From one strain of the psychometric approach, we borrow the notion that the control of attention is important (Engle *et al.*, 1999a,b). We differ from this psychometric approach in focusing on basic parameters that have difficulty-insensitive measures and are related to our own simple modeling framework (illustrated in Fig. 1). Borrowing from another strain of the psychometric approach, we also emphasize the importance of processing rates (Kail & Salthouse, 1994; Towse *et al.*, 1998). However, we have questioned the claim that a central, global processing rate controls performance quality. Instead, we have found separate, uncorrelated processing rates for memory search versus phonological processing operations (Cowan *et al.*, 1998).

We do not deny the importance of controlled attention but have concentrated our research instead on examining the limit in the capacity of the focus of attention in chunks. This may prove to be closely related to controlled attention, in which case it could provide a better defined type of measure of attentional processing limits than has been used in the psychometric approach. However, if such a correlation is not found, we still would hold that understanding attentional capacity limits is important in order to carry out an analysis of the demands of any particular task. (See Halford, Wilson, & Phillips, 1998, for theoretical work on how processing limits can account for developmental limits in the performance of various cognitive tasks.)

An open question in the psychometric subfield is the extent to which working memory processes are *modular* i.e., divided into separate components for separate domains of stimuli. According to some views, one process or mechanism may account for all that is important and interesting in working memory: most notably, the

quality of controlled attention (Engle *et al.*, 1999a) or general processing capacity (Daneman & Merikle, 1996; Just & Carpenter, 1992). According to other views, distinct pools of capacity exist for specific areas of processing such as linguistic processing (Waters & Caplan, 1996) or spatial processing (Shah & Miyake, 1996). The present approach is an intermediate one. We propose only one processing capacity limit, namely the limit in the focus of attention, which is general across modalities; but processing within activated memory may be susceptible to interference that depends on the similarity between stimuli, thus allowing for apparently modality-specific and materials-specific resources. One must know about the decay, interference properties, and search rates, not just attentional capacity, in order to know how well tasks can be carried out (Cowan, 1995, 1999a).

E. DIFFICULTY-INSENSITIVE MEASURES OF DEVELOPMENTAL CHANGE

1. Overview

A well-known, key epistemological problem that must be confronted in the field of cognitive psychology is that one cannot be sure how an observed measure (e.g., a pattern of proportions correct or reaction times across conditions of an experiment) can be interpreted to provide an estimate of an underlying factor of theoretical interest. This problem is compounded in the field of cognitive development, where one must find estimates of the factor that can be compared fairly across age groups. In the present case the problem is how to estimate a capacity limit, a search rate limit, or a decay rate (three parameters of special interest according to Cowan, 1995, 1999) from data collected across list lengths or retention intervals in immediate-memory experiments. We maintain that it is both possible and useful to derive absolute values for theoretically fundamental parameters at each age. This is in contrast to the usual approach (which we term a *relative-measurement approach*) in which the research focus is on relative differences between age groups on relevant tasks without any indication of the underlying, absolute values of parameters producing these age effects.

Potential concerns with this usual, relative-measurement approach are twofold. The first is a psychometric one: the results must be expressed in a manner such that the conclusions are not specific to the performance level of the individual or group. For example, suppose that decay of memory is observed across retention intervals and one wishes to determine if age groups differ in the amount of decay. The magnitude of decrement across retention intervals is likely to depend on the initial level of performance at the short retention interval. If performance is good initially, this well-learned memory may be more resistant to decay than a memory representation that is poorer initially. If older children perform at a higher level at short retention intervals than younger children do, an age difference in decay across retention intervals is not interpretable (because it could be an artifact of a stronger initial memory in the older children rather than a true decay

difference). A common solution to this sort of problem is to adjust the level of difficulty of the stimuli individually until each participant (or each group) is similar in the proportion correct at a short retention interval and then to observe decay (e.g., Keller & Cowan, 1994). More generally, the solution to such problems in developmental comparisons has often been to adjust the stimuli until the main effect of age disappears to determine whether there remains an Age \times Condition interaction that could be taken to indicate age differences in the process of interest (e.g., Massaro & Burke, 1991).

Given that type of solution to the measurement problem, however, one may wonder if the differences in stimuli across age groups make the developmental comparison difficult to interpret theoretically. In order to get the same level of performance at all ages, the adjustments in stimuli may lead to different sorts of limitations in performance for different age groups. For example, in a task requiring the comparison of two tones with a variable intertone interval, suppose for the sake of argument that performance can be based on either of two representations: a precise representation that decays quickly or a less precise representation that decays more slowly (e.g., sensory and categorical representations). In order to achieve equivalent performance at a particular, short intertone interval, the older participants have to be given tones that are less discriminable (Keller & Cowan, 1994). We do not know why this is the case but perhaps it is simply because younger children make more random errors, thus requiring easier comparisons to reach the same average level of performance. The stimulus adjustment might force older participants to rely most often on the more precise, unstable representation in order to hear very subtle differences between tones, whereas younger children's performance could be based more often on the less precise but more stable type of representation. That discrepancy in processing modes would lead to an underestimation of the stability of the memory representations in the older children relative to the younger ones.

The second concern with the relative-measurement approach is that it may be difficult to construct a defensible line of argument as to how a behavioral measure can be interpreted theoretically. The question is whether the observed differences across conditions reflect the process of interest or whether they instead reflect the intervention of some other, nuisance variable of lesser interest. For example, if one is interested in the duration of a sensory form of memory and its development, one wants to be able to know whether an observed age difference in forgetting over time is indeed due to differential sensory memory forgetting or whether it might instead reflect the better use of rehearsal, or better sustained attention across retention intervals, in older children; these being nuisance variables in the context of an investigation of sensory memory stability. Although many previous investigators have addressed these sorts of problems (e.g., Chapman & Chapman, 1978; Faust, Balota, Spieler, & Ferraro, 1999; Salthouse & Kausler, 1985), we believe that the field can still use additional, practical suggestions. In this regard, we make use

of various findings from the cognitive literature and point out that solutions to the psychometric and theoretical problems may converge. If one finds a means to identify the theoretical factor of interest, it should be possible to derive that factor similarly from results based on more than one level of difficulty of the stimuli.

Thus, the existence of a derived result that is stable across levels of difficulty of the stimuli can lend support to the argument that the result is theoretically meaningful and that it can be compared validly across age groups, for the same stimuli. In order to highlight such results and the methods underlying them, we suggest the term *difficulty-insensitive measures*. These, simply put, are measures that yield the same result across varying levels of task difficulty.

The advantage of difficulty-insensitive measures for developmental research is that they estimate the absolute values of parameters of information processing. They can yield conclusions about developmental change that are specific in magnitude and that do not depend upon the level of stimulus difficulty. This is in contrast to the usual tool of a group \times condition interaction, which typically varies in magnitude depending on the difficulty level of stimuli.

Few difficulty-insensitive measures have been found in developmental research; but there has not been much of a concerted search for these measures. In cognitive psychology, at least, two such exemplary parameters are rather well known. The *whole-report limit* described by Sperling (1960) is a limit in how many items can be recalled from an array (about four) which does not vary with the size of the array. It serves as an estimate of the capacity limit of working memory that has been confirmed in a variety of other work (Cowan, 2001a). The *memory search slope* described by Sternberg (1966) occurs in a procedure in which a probe is to be judged present in or absent from a recently presented set of items in memory and represents the slope of a linear increase in reaction time that occurs as a function of the number of items in the set held in memory. This slope reflects the amount of time it takes, per item, to search through the mental representations of items and does not vary with the exact list lengths presented, provided that the participant can retain the list. For example, the same slope would be obtained in experiments using List Lengths 1–4 or using List Lengths 3–6.

For both of the examples mentioned, there are remaining questions about exactly what processes these constant parameter values reflect (i.e., capacity of a visual store versus capacity of the focus of attention; serial versus parallel search of memory). Nevertheless, the difficulty-insensitive measures provide windows onto important processes that can be compared across individuals. Even without knowing all of the details of processing, an individual's profile of parameter values could be described. Also, when more is understood about the interpretation of the parameter values, more will consequently be known about the individuals whose values have been measured.

What essential properties of a measure would tend to make it difficulty-insensitive? One possibility is that the underlying parameter reflects a key

bottleneck in performance. Thus, in the procedure of Sperling (1960), performance is limited by some constant amount of output, presumably reflecting a constant limit in how much can be held simultaneously in working memory. Another possibility is that the parameter reflects a fixed process that can be used repeatedly on a trial, depending on the stimulus set. Thus, in the procedure of Sternberg (1966), it may take a certain, fixed amount of time to search for each item in a list, with the number of covert searches to take place on a trial dependent on the list length. The studies used as examples here will be explained more fully within the theoretical framework illustrated in Fig. 1, which included three processing mechanisms: *a capacity limit*, presumably in the contents of the focus of attention; *a retrieval rate*, presumably in bringing material from activated memory into the focus of attention; and *a decay rate* for activated information outside of attention (in the case of our studies, sensory information; though we assume that categorical information also can be in this activated memory). As mentioned above, these three mechanisms are central to the authors' view of working memory and therefore have been the basis of a great deal of our developmental research.

2. *Storage Capacity of the Focus of Attention*

The first parameter of processing to be considered is the capacity of the focus of attention. We have examined it using both attended but briefly presented visual arrays and ignored, spoken lists. In neither case can attention be used to group items into higher level chunks.

a. Visual Arrays. In one key study, Sperling (1960) presented brief spatial arrays of characters for partial or whole report and found that, although partial report (if cued quickly enough) allowed almost all of the items in the array to be recalled, the whole-report limit was about 4 characters. This whole-report limit remained fixed as the number of characters in the array ranged from 4 to 12. Initial interpretations focused on the possibility that the rate of extracting information from the sensory memory of the array, during the few hundred milliseconds while the sensory memory lasted, was the limiting factor for whole report. However, this interpretation is at odds with the finding that a much longer extraction period (several seconds) leads to a comparable whole-report limit in the auditory modality (Darwin, Turvey, & Crowder, 1972) and the finding that, in a procedure in which only one decision must be made on every trial (Luck & Vogel, 1997), a very similar recall limit of about 3.5 items can be derived (Cowan, 2001b; Cowan, Sauls, & Frisoe, in preparation). Cowan (2001a) and Cowan, Nugent, Elliott, Ponomarev, and Sauls (1999) interpreted these whole-report limits as reflecting a capacity limit coming into play when sensory information must be transferred to the focus of attention in order to allow the creation of categorical codes for the characters in the array, before they can be recalled. The amount of information that can be included in the focus of attention may be limited. This, in turn, places an identical

limit on how many items can be recalled, regardless of the size of the array (at least, within the range of 4–12 items).

The unit for this capacity limit presumably is the *chunk* of information, where a chunk comprises a group of items that, for one reason or another, has close mutual associations and is only weakly associated with other items in the stimulus field (Miller, 1956; Simon, 1974). For briefly flashed arrays like the ones Sperling (1960) used, grouping processes are limited and the chunk size is assumed to be the individual item (Cowan, 2001a). Thus, the whole-report limit can be viewed as a difficulty-insensitive estimate of the capacity of working memory in chunks. In adults, the limit is about four chunks.

Cowan (2001a) proposed that what is critical for the observation of a difficulty-insensitive measure of memory capacity is that the items presented are highly familiar and unassociated (and hence are individual chunks) and cannot be subjected to memorization processes that result in larger groups or chunks being formed at the time of presentation. Without such memorization processes, each item remains an isolated chunk in memory, so that the number of items recalled reflects the number of chunks recalled. Presumably, difficulty-insensitive measures of capacity can be obtained across memory set sizes (list or array sizes) because the capacity is limited to a fixed number of chunks.

Convergent findings come from a wide variety of procedures. One contribution of Cowan (2001a) was to organize findings into a taxonomy that helped to make clear the situations in which the presented items are single chunks in memory and, therefore, in which capacity limits (in terms of chunks recalled) could be estimated (in terms of items recalled). Four such situations were identified. First, in some situations the amount of information presented at one time overloads attention so that grouping cannot be carried out. The procedures of Sperling (1960), Luck and Vogel (1997), and Cowan *et al.* (1999) are good examples of this. Second, in some situations grouping processes are prevented directly, forcing the participant to rely on each item as a separate chunk and to hold the chunks in attention rather than a passive source. Studies in which rehearsal is discouraged (e.g., Waugh & Norman, 1965) or blocked (e.g., Murray, 1968) or in which the materials are not rehearsable (e.g., Glanzer & Razel, 1974) are of this type. Third, performance discontinuities often occur when the materials begin to exceed attentional capacity. The best known such discontinuity is the finding that when an array of objects is presented and the participant must report how many there are, rapid subitization is possible for up to about four items in parallel, whereas items over four must be counted serially (Mandler & Shebo, 1982). Other examples include (a) the ability to track up to about four objects moving through the visual field in different directions at once (Pylyshyn *et al.*, 1994); and (b) the absence of proactive interference in adults only for sets of up to four items presented at once, but proactive interference for larger sets (Halford, Maybery, & Bain, 1988). Fourth, and finally, there are indirect effects of capacity limits. Recall from large categories occurs in bursts

of up to about four items at a time, both in immediate recall (e.g., Ryan, 1969) and in long-term recall (Broadbent, 1975; estimated three, whereas Grassler & Mandler, 1978, estimated five). Even when experts can recall many more items than normal, the recall protocol shows answers coming in bursts of about four at a time (e.g., Ericsson, Chase, & Faloon, 1980). Also, mathematical models of cognitive processes typically use a working memory with a limit of four slots, presumably because such a model works well (Kintsch & van Dijk, 1978; Raaijmakers & Shiffrin, 1981). A reading of Cowan (2001a) and the following commentaries offers many potential reasons why a capacity limit of about four items works well, though we do not yet know which of these reasons have merit.

h. Spoken Lists. Some investigators have presented present auditory arrays that probably were too complex to allow rehearsal or chunking, a condition that allows capacity to be estimated (Darwin *et al.*, 1972; Rostron, 1974; Treisman & Rostron, 1972). Cowan *et al.* (1999) extended the logic further by showing how capacity limits could be estimated even for spoken lists in which only one item was presented at a time. To prevent chunking, a demanding distracting task was developed that did not create any acoustic interference (matching pictures with names that rhymed). When this distracting task was played during the presentation of lists of random, spoken digits and no response was required to most such lists, recall of the spoken lists that were occasionally probed were limited, in adults, to about 3.5 digits recalled in the correct serial position. This estimate, which is very similar to the whole-report limit described above, was obtained with lists of varying length. Presumably, all digits were automatically active in memory throughout the experiment but the binding between a digit and its serial position within the list occupied a slot as an "item" or chunk in the attentional, capacity-limited store. (This binding presumably occurs only if attention is turned to the sensory memory of the last list before it fades away.) Thus, in this method, attention was turned away from the auditory stimulus during its presentation to prevent grouping so that we could observe the limits in attentional capacity if attention was applied to the stimulus only at one point in time, when recall was probed.

3. *Rate of Retrieval into the Focus of Attention*

The second parameter of processing, the rate of retrieval of information into the focus of attention, has been investigated using both probed recall and the timing of serial recall.

a. Probed Recall. Another key study in cognitive psychology was that of Sternberg (1966, 1969). Given a target list to be remembered on every trial, the task was to indicate as quickly as possible, via a button-press, whether a probe item was in the list. Mean reaction times increased linearly as a function of the number of items in the list. The slope of the reaction time function across list

lengths was taken to indicate how long it took to search each item in the list to find an item identical to the probe. Sternberg considered the search process to be serial but subsequent investigators have suggested that the search actually could be parallel (Ratcliff, 1978; Townsend, 1976). In any case, the linear search function is not in question. The slope of the search function is the same when one increases from one to two items, two to three, three to four, and so on up to six items. In that sense, the slope is a difficulty-insensitive measure of the rate of mental search through the list.

b. Pauses in Spoken Recall. Cowan (1992) examined memory span in 4-year-old children and measured the duration of several segments of responses for errorless trials: (a) the silent preparatory interval between the end of the stimulus list and the beginning of the response, (b) the duration of each word in the response, and (c) the duration of each silent interword pause between words in the response (recorded as 0 if there was no such pause). The mean interword pause increased as a function of the list length, which suggested that children searched through the entire list to select the item to be spoken next. The longer the list was, the longer the search took. Subsequent research established the same phenomenon for older children (Cowan, 1999b; Cowan *et al.*, 1994, 1998) and adults (Hulme, Newton, Cowan, Stuart, & Brown, 1999; Tehan & Lalor, 2000).

Other aspects of the findings show further similarities between the probe reaction time and interword pause time measures. In particular, neither is affected by the duration of stimulus words (Cowan *et al.*, 1994; Chase, 1977; Clifton & Tash, 1973), suggesting that an abstract node is searched, not a phonological representation as in the word length effect in serial recall (Badddeley, 1986; Badddeley, Thomson, & Buchanan, 1975). Neither shows primacy effects, or enhanced recall of the first few list items, although the probe reaction time measure sometimes shows recency effects, or enhanced recall of the last few items (e.g., McElree & Doshier, 1989; Monsell, 1978; Ratcliff, 1978). Moreover, at least in children's recall of up to about five items, interword pauses are flat across serial positions except for a slight shortening at the last serial position (Cowan, 1992; Cowan *et al.*, 1998). If participants were able to eliminate each item from the search process after it was recalled, the pauses should have become steadily shorter across serial positions of the recall process, but that did not appear to occur. For longer lists in adults, longer pauses probably occur between some items in the response because the items are grouped together (Anderson & Matessa, 1997), albeit group boundaries may occur at different locations for different individuals. Those intergroup boundaries cannot, it appears, be seen clearly within shorter lists.

Unpublished evidence from the data set of Cowan *et al.* (1998) show that the interword pauses form a difficulty-insensitive measure of the underlying search process. Most children in the first, third, and fifth grades had pauses for lists of two, three, and four spoken digits and the pauses show a nearly perfect linear increase

within that range in all groups. That is to say, the increase in the pauses between two- and three-digit lists is the same as the increase between three- and four-digit lists. This difference represents the amount of time to search for a single extra item in the list, and this per-item search time is a difficulty-insensitive measure of the retrieval rate.

4. *Sensory-Memory Decay Rate*

The third parameter of processing to be described is the decay rate of one type of automatically activated memory, namely sensory memory. A difficulty-insensitive measure of sensory memory cannot yet be identified with confidence equivalent to what we have suggested for capacity and retrieval speed. However, a possible measure is inspired by the early studies of sensory memory, conducted not only by Sperling (1960) and Darwin *et al.* (1972) but by many others (for reviews see Broadbent, 1958; Coltheart, 1984; Cowan, 1984, 1988, 1995; Crowder, 1976; Di Lollo & Dixon, 1988; Massaro & Loftus, 1996; Nairne 1990; Penney, 1989). The literature shows that we have a rich and vivid, but short-lived, memory of the sensation arriving in any particular modality, which tends to be overwritten by subsequent stimulation in the same modality (especially if physical features of the subsequent stimulus are similar to those of the to-be-remembered stimulus). The memory for sensation can be divided into two types, presumably in any modality: a brief mental afterimage that lasts only 1/5th to 1/3rd of a second and is perceived as a continuation of sensation and a longer memory for sensation that is perceived instead as a vivid recollection for up to about 20 s (Cowan, 1988, 1995). To examine the latter, longer memory without encountering sensory interference, one needs to test performance on the last stimulus in a list. The list length must be varied to demonstrate that the memory is difficulty-insensitive (i.e., yielding the same measurements regardless of the list length). If the stimuli can be retained in memory as labels, one must provide a powerful distracting task to prevent such labels from being formed so that the participant can only rely on sensory memory as opposed to memory for categorical labels; yet the distracting task should be in another modality so as not to cause sensory interference.

Perhaps no extant data meet all of these requirements. However, some relevant data come from a developmental study by Cowan, Nugent, Elliott, and Saults (2000). In that study, attention was directed toward a silent task in which pictures were matched on the basis of rhymes while random lists of spoken digits, presented at a 2/s rate, were ignored. Occasionally, a change in the computer screen display indicated that the participant should use the number keypad to recall the digits in the most recent spoken list. The retention interval between the end of the target list and the recall cue, filled with a continuation of the rhyming game, lasted 1, 5, or 10 s. The list was always as long as the longest list that the participant had recalled without error in a previous task, in which he or she had paid attention to the digits (which we refer to as the participant's *maximal span*). We thus have

ignored-digit recall data for only one list length per participant, namely the participant's maximal span.

Although difficulty insensitivity can only be examined properly with multiple difficulty levels in the same participant, we can show that a measure of decay, the proportion decrement in the last serial position over retention intervals, is at least not very sensitive to performance level differences within an age group. Sufficient data to examine this exist only among the youngest group in the study, the second-graders (those with a span of 4, $N = 7$; span of 5, $N = 8$; and span of 6, $N = 6$). Among these children, after a 1-s retention interval, the proportion correct at the last serial position depended heavily on the span: it was .82, .63, and .46 for children with a span of 4, 5, and 6, respectively. Yet, the ratio of the score after a 5-s retention interval to the 1-s score (indicating the proportion of memory remaining after 5 s) was similar in the three span groups: 0.52, 0.55, and 0.55 for children with a span of 4, 5, and 6, respectively. Thus, regardless of span among these children, the half-life of memory for the last list item in this memory-for-ignored-speech task was about 4 ($= 5 - 1$) s.

5. *Application of Difficulty-Insensitive Measures to Development*

In sum, so far we have seen that certain measures remain fixed as one changes the array size or list length: the number of items that can be recalled from a complex array or an unattended list, the rate of retrieval of information in the absence (we suspect) the rate of decay of unattended acoustic information in the absence of acoustic interference. Having such measures is important for cognitive research because they provide stable estimates of processing parameters that are not highly dependent on the particular difficulty of materials. They are doubly important for developmental comparisons because different age groups show different overall levels of performance (see also Cowan, 2000; Cowan, Saults, Nugent, & Elliott, 1999). With difficulty-insensitive measures, developmental comparisons of data patterns are possible that are both (a) measured at the same overall performance level across age groups and, simultaneously, (b) measured using the same stimulus sets across age groups. That possibility, if extended and verified, represents an exciting opportunity for developmental research. It leads to the possibility of estimates of the quantity of change in basic processing mechanisms with development.

II. Evidence of Developmental Change in Processing Parameters of Working Memory

Having discussed the theoretical framework, including three parameters of working memory and how they can be measured, we now describe evidence of their development.

A. CAPACITY LIMITS

1. Preliminary Evidence from a Spatial Array Task

Working memory performance, even on the simplest tasks, clearly increases with age in childhood. However, why that change takes place is not clear. Researchers have been equivocal about whether there is a change in the capacity of working memory (e.g., Case, 1995; Pascual-Leone, 1970) or just a change in how well a fixed capacity is used (e.g., Case *et al.*, 1982; Karl, 1990). To address this issue, we examined capacity in developmental studies using measures for which the number of chunks can be estimated, inasmuch as each stimulus is a familiar chunk and the conditions discourage additional chunking, as explained by Cowan (2001a).

In an unpublished study (Cowan, Fritsioe, Elliott, Saults, Brunner, & Lacey, in preparation), we looked for evidence that the type of working memory needed to retain an array of color squares, like that examined in adults by Luck and Vogel (1997), may change with age in childhood. The method was nearly identical to that of Luck and Vogel. Each trial consisted of an array of 4, 6, 8, or 10 randomly placed color squares and then another array that was identical to the first or differed in the color of one square. A single cue (a white circle) surrounded one square in the second array. If a color had changed, this cue surrounded the changed color. The task was just to indicate whether that square had changed color between the two arrays or had remained unchanged. Any particular color could occur more than once in a single array so the participant had to encode not just the presence or absence of each color, but also the binding between the color and its spatial location in the array. In theoretical terms we assume that all of the colors become more or less activated in memory across trials but that the binding between colors and locations is subject to a more severe capacity limit.

We evaluated the results with a variation on a simple formula devised by Pashler (1988). To estimate an individual's memory capacity k , we assumed that out of N items in the first array on a trial, k of them were present in working memory. Thus, as a result, the right answer would be known on k/N of the trials. On the remaining $(N-k)/N$ of the trials, we assumed that the participant guesses "yes, there was a change" with a probability g . Then it follows that hits $H = k/N + [(N-k)/N]g$ and correct rejections $CR = k/N + [(N-k)/N](1-g)$, the latter differing from Pashler. Combining these equations, it can be shown that $k = N(H + CR - 1)$. This rough-and-ready estimate appears to have some validity in that it resulted in constant estimates of the capacity limit across array sizes, provided that the array size was larger than the capacity (which is necessary to avoid ceiling effects in the more capable participants, i.e., compression of the range of results).

For 56 adults the capacity estimates were, for set sizes 4, 6, 8, and 10, respectively, 3.15, 3.77, 3.84, and 3.80. For 61 fourth-grade children they were 2.28, 2.26, 2.39, and 2.36. These means are rather stable across set sizes, except for set size 4 in adults. Given that some adults actually had a capacity below 4 and others

had a capacity above 4, this small set size leads to a ceiling effect. Otherwise, the method is a difficulty-insensitive measure of capacity.

2. Evidence from a Serial Recall Task

a. Introduction. At least since the article by Miller (1956), the limit in short-term memory capacity has been considered a major aspect of human thought. However, the "magical number seven" that Miller observed results from a complex processing system in which stimuli are grouped together, rehearsed, and partly memorized to the point that one cannot tell how many independent ideas or "chunks" are concurrently occupying the short-term memory faculty. Obtaining measures of the number of independent chunks has been difficult, for example, in short-term serial recall. Cowan, Nugent *et al.* (1999a) investigated this issue developmentally with an ignored-speech procedure modeled after that of Saults and Cowan (1996), but with only one retention interval and with each participant receiving ignored lists of four different lengths. The rationale was that one can prevent rehearsal and memorization by presenting items one at a time and diverting attention away from them. This seemed better for developmental research (because it is simpler) than the usual means of preventing these processes, which has been by presenting items in a complex spatial or temporospatial array (Darwin *et al.*, 1972; Luck & Vogel, 1997; Sperling, 1960). In both cases, though, an indication that one has succeeded in preventing the formation of higher order chunks is that the number correct does not change with the size of the array or list and matches other estimates of capacity (Cowan, 2001a). Given that pattern of responding we assume that, from the large amount of information in sensory memory, the independent, unrehearsed items that can be drawn into the short-term memory faculty (presumably, into the focus of attention) are limited to a fixed number.

b. Basic Method of Cowan et al. (1999a). To investigate capacity using ignored speech we tested 72 participants in all: 24 in each of 3 age ranges (Grade 1, Grade 4, and adult). After elaborate training and familiarization procedures, each participant carried out an auditory memory task and a visual distracting task, both separately and jointly. In the distracting task, some standard names of various pictures of common nouns were learned in advance and practice was given. Each display consisted of four pictures at the corners of the computer display (e.g., *tail, boy, hat, and rock*) and a central picture whose name rhymed with one of the peripheral pictures (e.g., *mat*). The task was to use the computer mouse to select, as quickly as possible, the peripheral picture with a name that rhymed with that of the central picture (e.g., *hat*). Upon completion of that response, the central picture changed and the participant was to answer again. Thus, the pace of the task was self-adjusting. As we explain shortly, this primary task was reasonably effective in keeping attention away from the spoken words during their presentation.

Before each of these tasks, each participant received an aurally presented span test with a spoken response. Then, in the auditory-memory-alone task (serving as an "attended-speech control" for a subsequent, ignored-speech condition), lists of spoken digits were presented at a rate of 0.5 s/digit and were to be recalled in the presented order using the numerical keypad. The developmental difference was no larger with a keypad response than with a spoken response in the span test, and therefore we concluded that the keypad played no role in the developmental findings. (The same was true in Cowan, Nugent *et al.*, 2000.) In the auditory-memory-alone task, lists were presented through headphones at four lengths: the longest length that the participant had repeated correctly in the span task (which we term "span length") and the three adjacent, shorter lengths (span-1, span-2, and span-3). A child managing to repeat one five-digit list correctly in the span test, for example, would receive five-, four-, three-, and two-digit lists in random order in the auditory memory tests.

The visual task was carried out alone in two test segments and also, most critically, in a dual task while spoken lists of digits were to be ignored. These spoken lists were presented at a fast, two-per-second rate and occurred with randomly mixed silent periods of 1, 4, or 7 s between lists. A to-be-ignored series continued for 45 to 100 s and then the visual task was interrupted with a response screen indicating that the participant was to use the keypad to recall the last spoken list, which had ended 2 s previously. The response screen included one empty box for each of the numbers that had occurred in the list, so the participant knew the list length. After the participant responded, a new set of pictures appeared and the primary task resumed. After the dual task session, the visual-alone task and then the auditory-memory-alone task were presented again.

c. Adequacy of the Distractor Task. Cowan, Nugent *et al.* (1999a) showed that memory for ignored speech yielded estimates of capacity that are constant across list lengths, yet change with age, as explained further below. The constancy of estimates across list lengths presumably was obtained because recall was based on information transferred, just after the response screen appeared, from a large-capacity sensory memory representation of the list to a limited-capacity focus of attention. There was no prior opportunity to convert the sensory memory to larger chunks of information before recall, so each retrieved item remained a separate chunk in the focus of attention and the number of items that could be retrieved therefore was determined by the capacity limit. That is the same interpretation given by Cowan (2001a) to procedures in which information must be gleaned from a rapidly presented, simultaneous array of characters, again producing a capacity estimate that is constant across set sizes (e.g., Sperling, 1960).

An alternative interpretation of the ignored-speech memory results, however, is that there actually is no memory for "ignored" speech per se and that the memory that we observed depended upon attention diverted from the primary, visual task. Such an interpretation would invalidate our measure of capacity. However,

countering this interpretation are several strong arguments that depend largely on data collected in three developmental studies with the rhyming task (Cowan, Nugent *et al.*, 1999a, 2000; Sauls & Cowan, 1996). The patterns of performance for attended-speech and unattended speech trials differ considerably and the ignored-speech memory capacity estimate converges with results from very different, and varied, methods of estimating capacity (Cowan, 2001a). As a further argument against the attention-diversion hypothesis, shifts of attention to the speech channel, though they may occur occasionally, cannot account for several critical features of the results that we have obtained:

1. The auditory stimuli are highly repetitive, irregular, and presented every few seconds but cued for recall only every minute or so. The subjective impression is that it is almost impossible to train attention on the channel to be ignored while performing the visual task, just as one would expect from habituation of the orienting response (e.g., Sokolov, 1963; Cowan, 1988).
2. In all three studies, we have recorded the primary task reaction times and accuracies, and have included *primary-task-alone* control test phases just before and just after the ignored-speech session. These measures invariably show that, for each age group, performance in the initial primary-task-alone phase is nearly identical to primary task performance during the immediately following, first block of ignored speech; and although performance improves during the ignored-speech task, primary task performance during the last block of ignored speech is nearly identical to primary task performance in the immediately following, primary-task-alone phase. Thus, speech does not appear to shift attention away from the primary task. Reaction times decrease across age groups, as expected, but visual trials were self-paced and thus occurred commensurately faster in the older age groups.
3. In two studies, to test for possible trade-offs between tasks, we have examined within-age correlations between mean performance on the primary task and mean memory for ignored speech. These correlations are very small (and nonsignificant) for children and positive for adults. Neither result is consistent with the suggestion that participants trade off performance on the two tasks, which should produce a negative correlation. Similarly, the correlation between the primary task and memory performance on a trial-by-trial basis *within* an individual, in memory for ignored spoken lists (Cowan, Nugent *et al.*, 2000), is near zero for first- and fourth-grade children and is significantly different from zero in adults, but in a direction in which trials with better primary task performance also go along with better auditory memory. Perhaps adults in a more alert state can do both tasks better because alertness facilitates the process of retrieving information from sensory memory. At any rate, there is no evidence of a trade-off between tasks.
4. Finally, in all age groups, the pattern of forgetting (Cowan, Nugent *et al.*, 2000) was not what would be expected if the spoken digits were attended

Initial and final serial positions showed considerable forgetting across retention intervals. That differs from retention across filled intervals in *attended* speech tasks, in which the good performance in the initial serial positions (primacy effect) remains stable across retention intervals (Jahnke, 1968). Thus, the typical pattern caused by attention is absent for the ignored lists. Moreover, the *age difference* in forgetting across retention intervals is restricted to the final serial position of the list, the only one at which there are no subsequent sounds to interfere with sensory memory (Crowder & Morton, 1969). Covert strategies of monitoring the ignored channel would have led to the prediction of an age difference more extensively across the list positions (and a trade-off between tasks).

d. Results of the Memory Task of Cowan, Nugent et al. (1999a). In the method of scoring used for the auditory memory task, credit is given only for items recalled in the correct serial position. The underlying assumption is that all nine digits become activated in memory, given that the response set consists only of these same digits. What actually constitutes an "item" in this task, therefore, is the association between items and serial positions in the list. These associations, or "bindings," form an episodic record of the trial. The number correct is taken to reflect the capacity of the focus of attention.

The pattern of responding in the memory tasks is shown in Fig. 3. The dashed lines indicate memory for attended speech, which increased across list lengths

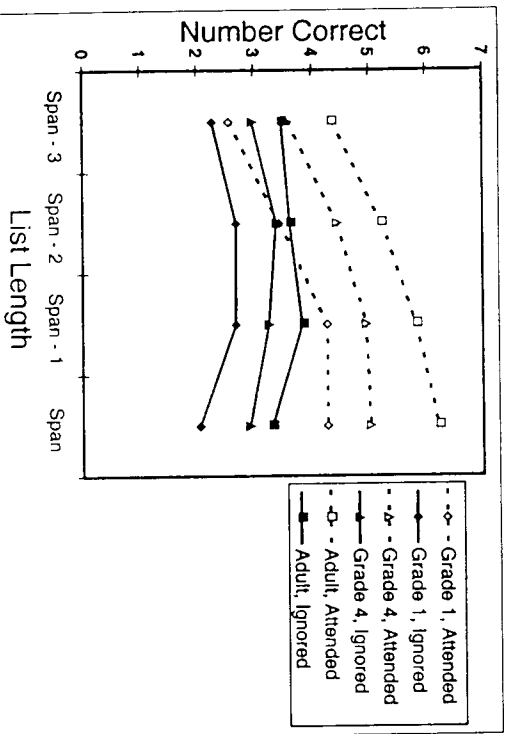


Fig. 3. Mean number of correctly recalled digits from attended lists (dashed lines) and unattended lists (solid lines) as a function of relative list length for participants in three age groups (graph parameter). Notice that the number correct is fixed across list lengths for the unattended lists, indicating a capacity limit that increases with age. Adapted from Cowan *et al.* (1999a).

(x axis) for all age groups (graph parameter). In contrast, the solid lines indicate memory for ignored speech, which was rather constant across list lengths. The age differences in the levels of these solid lines, which were significant, indicate that the capacity of memory increased with age. Also, the fact that the age differences in the attended-speech control condition were about the same size as age differences in the ignored-speech condition (true especially between Grades 1 and 4) suggests that age differences in span do not stem mainly from attention-demanding processes taking place during the reception of the list (e.g., rehearsal), in contrast to what is commonly assumed. (Rehearsal accounts were questioned also by Cohen *et al.*, 1985; Dempster, 1981; and Huttenlocher & Burke, 1976.) If attention during reception had played a major role, then age differences should have been substantially larger in the attended- than in the ignored-speech condition.

e. A Difficulty-Insensitive Measure of Memory Capacity for Spoken Digits.

The data shown in Fig. 3 illustrate that performance, measured as the number of items recalled in the correct serial position, was not affected much by the number of items in the list. In that figure, however, as in the experiment carried out by Cowan, Nugent *et al.* (1999a), the range of list lengths was adjusted according to the participant's ability in the attended-speech condition. Cowan, Nugent *et al.* also found a similar pattern when the data in the ignored-speech condition were categorized according to the absolute list length. Lists of four, five, and six items appeared in the data sets of most participants. A plot of the number correct for these list lengths, and for span-length lists, reveals that the developmental trend was approximately the same regardless of the list length (Fig. 4). This difficulty-insensitive developmental trend strengthens the evidence that the capacity changes across ages and

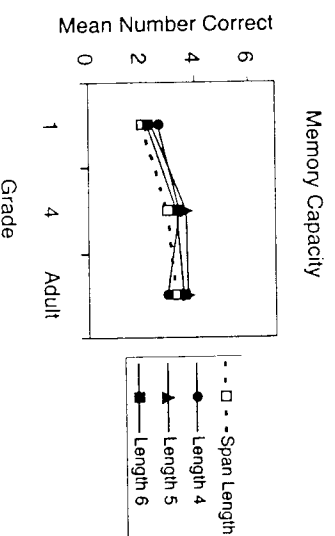


Fig. 4. Mean number of correctly recalled digits from unattended lists as a function of age for several list lengths (graph parameter) from the same data set as in Fig. 3. The irrelevance of list length for the developmental function shows that this is a difficulty-insensitive measure of capacity. After N. Cowan, J. S. Sauls, L. D. Nugent, and E. M. Elliott (1999b). The microanalysis of memory span and its development in childhood. *International Journal of Psychology*, 34, 353-358. Copyright 1999, International Union of Psychological Science.

does not simply appear that way because of some complex interaction between the stimulus conditions and ability levels at different ages. The capacity measure is presumably difficulty-insensitive because, for these unattended spoken digits, capacity is limited to a fixed number of digit-to-serial-position bindings regardless of the number of digits present in the list.

B. RETRIEVAL RATES

1. Introduction

The full understanding of performance in any memory task must take into account both the level and the rate (or timing) of performance. Yet, few studies of immediate recall have used timing measures because of the considerable tedium of measuring this timing. We have found it rewarding to do so. Cowan (1992) timed 4-year-old children's responses to correctly repeated lists, using a computer-based speech waveform editor to obtain the duration of each segment in recall in what may be the first such study of span recall timing. Neither the durations of *preparatory intervals* between the stimulus and response lists nor the durations of *words* in the responses provided a clear indication of the ability level of a child. However, durations of the *interword pauses* between words in the response did. As mentioned above, children with a higher span repeated lists of a particular length with shorter interword pauses in their responses. Cowan *et al.* (1994) found that a given list was repeated with longer interword pauses in 4-year-olds than in 8-year-olds.

The processes taking place during the pauses did not appear to include verbal rehearsal, inasmuch as the pauses were no different for lists of monosyllabic versus multisyllabic words (Cowan *et al.*, 1994), whereas covert verbal rehearsal would have been expected to take much longer for the multisyllabic words (Baddelley, 1986). However, the pauses within a particular individual's recall increased as a function of the list length and remained relatively constant across serial positions in the list. This suggested that the pauses reflect a sort of memory search process in which the entire list plays a role during each pause, during which the participant searches throughout the list for the word to be recalled next (not necessarily in a serial manner; for a parallel model see, for example, Ratcliff, 1978).

2. Findings of Cowan, Wood *et al.* (1998)

Cowan, Wood *et al.* (1998) reported results of a study of the timing of spoken recall in a digit span task in first-, third-, and fifth-grade elementary school children ($N = 24$ per age group). This study focused on correlations between span and timing measures and demonstrated that age effects do not necessarily show the same pattern as individual differences. To see this, one must distinguish between the preparatory interval in recall (the relatively long silent period between the end of a stimulus list and the beginning of the first word in the response) and the

interword pauses (the typically shorter periods between words in the response). Within correct recalls at a particular list length, older children had markedly shorter preparatory intervals. Nevertheless, preparatory intervals were completely unrelated to span across all 72 children. (The age differences were not large enough to force that correlation.) In contrast, the interword pauses both decreased with age and correlated with memory span.

Cowan, Wood *et al.* also examined children's ability to repeat digits rapidly. Baddelley (1986) and others have found that this type of measure correlates with memory span. Most previous measures of rapid-speaking ability have involved the repetition of a small set of items over and over. However, we worried that such a measure confounds the time needed for speaking with the time needed for planning each cycle of the repetition. Therefore, we used a rapid-speaking measure in which the children simply counted from 1 to 10 once on each trial as quickly as possible and a measure in which a short list of one to four randomly ordered digits was presented for repetition once on each trial following a start signal (a tone) so that the planning period is separable from the speaking period. The same list was presented for six trials in a row in order to examine the role of memory in the rapid-speaking task, but essentially the same correlation between speaking rate and memory span emerged in each of the six repetitions despite large practice effects across repetitions. Digits span was correlated with tasks of counting 1–10 and repeating two-, three-, and four-digit lists.

Various authors have suggested models in which memory development can be traced to a change in the global rate of processing, which in turn is assumed to influence the rate of processing of various specific skills (e.g., Kail & Park, 1994; Kail & Salthouse, 1994; Fry & Hale, 1996). If that were the whole story, then there should be substantial correlations between interword pause measures and rapid-speaking measures, given that part of each measure would reflect a common global rate of processing. However, Cowan, Wood *et al.* (1998) found no correlations between these two measures, which picked up different portions of the variance in digit span. (The mean of 12 correlations was .07.) Together in a latent variable model presented by Cowan, Wood *et al.*, in which the two types of processing rates both are influenced by age and both contribute to span independently, they accounted for 60% of the total variance in span and 87% of the age-related variance in span. This model contained three latent variables: short-term memory retrieval duration, rapid-speaking duration, and span. The path coefficient from age to short-term memory retrieval duration was $-.42$; from age to rapid-speaking duration, $-.30$; from short-term memory retrieval duration to span, $-.41$; and from rapid-speaking duration to span, $-.49$. These paths were significant but the path from age directly to span was nonsignificant (.17) within this model. Correlation between disturbances for the two latent variables for different kinds of rates was near zero, showing that they are unrelated. This model fit significantly better than a model with only one latent variable encompassing all of the rate measures. Thus, individuals

have different ability profiles, reflected in the two types of unrelated rate measures, that can result in equivalent memory spans.

The model was verified in a second experiment with 180 adults, using a different measure of rehearsal rate (covert repetition of the alphabet or the numbers 1–10 with the participant manually marking the beginning of each cycle of rehearsal) and memory-search tasks modeled after Sternberg (1966), in place of interword pauses, to measure short-term memory retrieval. These two types of measures correlated with span but not with each other and together picked up 30% of the adult variance. Thus, the pattern is not a fluke and has considerable conceptual validity.

In the simple model of Cowan, Wood *et al.* (1998), the two latent variables were identified as retrieval rate and rehearsal rate, concepts that seemed theoretically most reasonable at the time. However, we have subsequently questioned the popular notion (Baddeley, 1986) that the rapid-speaking rates reflect rehearsal. Various studies indicate that rehearsal develops late in the elementary school years (e.g., Flavell *et al.*, 1966; Henry, 1991; Ornstein & Naus, 1978). Cowan (1999b) carried out a reanalysis of the Cowan, Wood *et al.* (1998) data focusing on within-age patterns and found that the rapid-speaking rates correlated with memory spans within Grade 1, but not within Grades 3 or 5. This difference cannot be attributed to a greater sensitivity of correlations in Grade 1 because a very different pattern was obtained for interword pauses, which correlated with memory span in Grade 5, but not in the younger grades. The correlation between rapid-speaking rate and span at such a young age suggests that this rate reflects a verbal ability that is functional in first grade, such as the ability to plan or retain phonological materials (Gathercole & Hitch, 1993), rather than cumulative rehearsal.

The following analysis focuses on the durations of interword pauses and not on the durations of phonological processes. The evidence reviewed above provides assurance that these interword pauses reflect a specific retrieval process rather than a fully general speed of processing. As discussed above, the pauses increase with the list length but do not depend on the length of individual verbal stimuli and are uncorrelated with rapid-speaking ability.

Interword pauses that have been observed in various studies appear to be rather uniform across serial positions of the list (e.g., Cowan, Wood *et al.*, 1998). One might have expected that pauses would differ across serial positions if participants grouped the list items and repeated them with a timing indicative of the grouping boundaries (Anderson & Matessa, 1997). Perhaps that would be the case with adults, or with relatively long lists in children capable of recalling them correctly, although there still would be the problem of individual differences and list-specific differences in where group boundaries are placed. For the lists of digits that most of our children have recalled successfully, namely the two-, three-, and four-item lists, we have found no evidence of grouping. We have asked second-grade children about their recall methods and they make little or no mention of grouping, unlike adults, who usually report some type of grouping strategy. Therefore, the interword

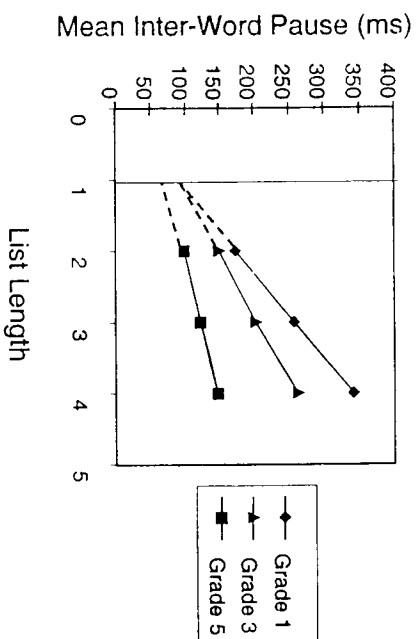


Fig. 5. Interword pauses as a function of list length for three age groups (graph parameter). From the data set of Cowan, Wood *et al.* (1998)

pauses for these list lengths appear to be relatively pure indices of a retrieval time with little contribution of grouping processes.

Figure 5 shows the interword pauses as a function of age and list length, for all children capable of correctly repeating lists of these lengths. This figure shows a very orderly pattern in which the y intercept of the performance function differs little with age but the slope of retrieval, as a function of list length, decreases systematically across age groups.

Figure 6 plots the same data in a different manner, with age group on the x axis and list length as the graph parameter. This allows the addition of a relative list

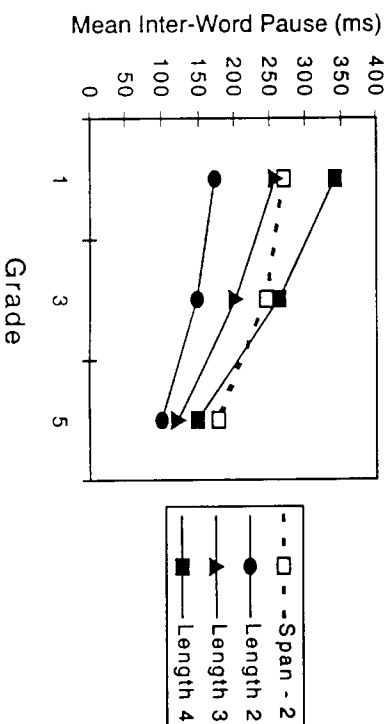


Fig. 6. Interword pauses as a function of age group for several list lengths (graph parameter), from the same data set as in Fig. 5. After N. Cowan, J. S. Saults, L. D. Nugent, & E. M. Elliott (1999b). The microanalysis of memory span and its development in childhood. *International Journal of Psychology*, 34, 353–358. Copyright 1999, International Union of Psychological Science.

length; only Length Span-2 included enough trials to form stable means. This figure shows that interword pauses per se are not a difficulty-insensitive measure inasmuch as they shorten across ages to an extent that strongly depends on the list length.

3. *Supplementary Results: A Difficulty-Insensitive Measure of Retrieval Rate*

Although interword pauses do not form a difficulty-insensitive measure, the results shown in Fig. 5 suggest that the slope of the search function is difficulty-insensitive. Thus, the amount of change in pause durations between two- and three-item lists, for any age group shown in the figure, is nearly identical to the amount of change between three- and four-item lists. That per-item increase or search slope can be interpreted as the speed of memory search, analogous to Sternberg (1966).

Given that the intercepts of the functions shown in Fig. 5 are small in absolute terms and differ little between groups, the search slope is by far the most important factor determining the duration of interword pauses. Consequently, one can get a nearly difficulty-insensitive measure by taking each mean pause duration shown in Fig. 6 and dividing it by the list length. The result is plotted in Fig. 7. Notice that, in contrast to Fig. 6, this per-item pause amount does approximate a difficulty-insensitive measure. The resulting measure changes with age in about the same way regardless of the list length. If the measure is taken as an estimate of the memory-search rate, the estimate clearly is similar to what has been found in studies of the development of memory search using probe-recognition procedures (e.g., Keating, Keniston, Manis, & Bobbit, 1980).

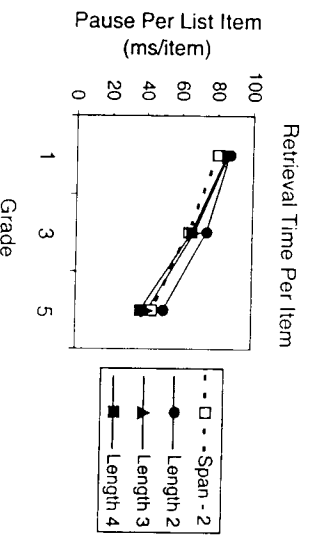


Fig. 7. Pause per list item, calculated by dividing the mean pause by the list length, for three age groups (graph parameter). From the same data set as Figs. 5 and 6. Notice that this derived measure is approximately difficulty-insensitive; the small, similar intercepts shown in Fig. 5 are necessary for this difficulty insensitivity to emerge. After N. Cowan, J. S. Sauls, L. D. Nugent, and E. M. Elliott (1996b). *The microanalysis of memory span and its development in childhood*. *International Journal of Psychology*, 34, 353-358. Copyright 1999, International Union of Psychological Science.

C. DECAY RATES

1. Introduction

According to the theoretical framework of Cowan (1988, 1995), various parameters of processing are interrelated. An example is the relation between rates of retrieval and decay. The relation depends on the description of short-term memory retrieval as a process in which information is drawn from an automatically activated source into the focus of attention. Given that the automatically activated memory representation of a stimulus is short-lived, the ability to retrieve information from that representation depends not only on the retrieval rate, but also on how long the activated representation lasts and how soon it is accessed.

2. Findings

Sauls and Cowan (1996) used an ignored-speech procedure to examine the forgetting function in children in Grades 1 and 3 and adults. Instead of digit lists, the stimuli were randomly presented repetitions of four isolated spoken words: *bee*, *red*, *bow*, and *roe*. (These four words were helpful because they permitted a separate assessment of memory for consonants versus vowels, though the results were similar for both types of speech sound.) The participant played a silent computer game while ignoring random repetitions of these words presented through headphones with 1, 5, or 10 s between words. Most ignored presentations were never tested. However, occasionally (every few minutes), the computer game would suddenly be replaced by a response screen consisting of pictures of the four spoken words. At that point the task was to select the picture corresponding to the spoken word that had been presented last, 1, 5, or 10 s ago (the retention interval), presumably based on the sensory memory trace that can be lost during this time between the spoken word presentation and the recall cue.

Cowan, Nugent *et al.* (2000) considered that the developmental increase in memory persistence observed by Sauls and Cowan (1996) might be attributed to differences in the effective memory load imposed by isolated words. For example, an isolated word is two items below span for a child with a span of three words, but four items below span for a child with a span of five words. We consequently adjusted the list lengths so that each child was tested with spoken lists of digits at a length equal to his or her own span (defined as the longest list that could be recalled, given that an integer value was needed). Participants were second- and fifth-grade children and adults. Lists were presented at a 0.5-s-per-item rate to maximize list continuity in sensory memory. The silent computer rhyming game of Sauls and Cowan (1996) was again used for the primary task in the ignored-speech trials. Attended-speech trials, administered both before and after the ignored-speech trials, showed no loss of information about the spoken lists across retention intervals and no age differences. For ignored speech, averaged across all serial positions,

the auditory memory data showed a steep forgetting function across retention intervals, but no significant age difference in this forgetting function.

An age difference was revealed when performance at each serial position was examined separately. The expectation from previous research was that sensory memory should be most clearly revealed at the end of a list (e.g., Balota & Engle, 1981) because the last serial position is the only one not followed by additional acoustic stimuli that can cause interference with sensory memory. These analyses showed no age difference in forgetting except, as expected, in the final serial position. To illustrate this pattern in a simplified manner, Fig. 8 shows memory for ignored speech across retention intervals for just three of the serial positions. Although for the two groups of children the first and last serial positions were at similar levels of performance after 1 s, these serial positions show very different patterns across retention intervals. This difference was absent at the first serial position (left panel), as in all of the medial serial positions (e.g., middle panel), but was very clear in the final serial position (right panel). At this position, as shown in the figure, the adults were above the children at all retention intervals so that their forgetting functions could not be adequately compared to the children. However, the data clearly show that forgetting of list-final items was more severe in the younger children than it was in the older children. This result is the same as Saults and Cowan (1996) found for isolated words. Again, the primacy effect shown for ignored lists (Fig. 8) differs from *attended* lists with a filled retention interval (Jahnke, 1968), which show substantial forgetting in the recency portion of the list but not in the primacy portion.

The findings lead to important points about age differences in memory. If those differences resulted primarily from the covert use of strategies during the presentation of lists that were supposed to be ignored, then these age differences would be expected to show up in the primacy portion of the serial position function, given the usual assumption that rehearsal strategies are used most at the beginning of a list (e.g., Atkinson & Shiffrin, 1968; Rundus, 1971). The occurrence of the age difference in forgetting in the final serial position only (see Fig. 8) appears to confirm the suggestion that what develops here is the persistence of auditory sensory memory or some other form of passively held memory activation. (Primacy effects in all age groups at a short retention interval can be attributed to the distinctiveness of items at either end of the list; see, for example, Naime *et al.*, 1997.)

We have not yet conducted a study to examine directly whether the age difference in decay rate is difficulty-insensitive, although preliminary indications are encouraging. What is necessary to design a difficulty-insensitive measure is to formulate an objective rule that makes sense and that yields the same estimates of the magnitude of age-related change across a range of difficulty levels (i.e., in our studies, list lengths or set sizes). In the case of decay rate, the rule is to examine the proportion correct at the final serial position. If performance at

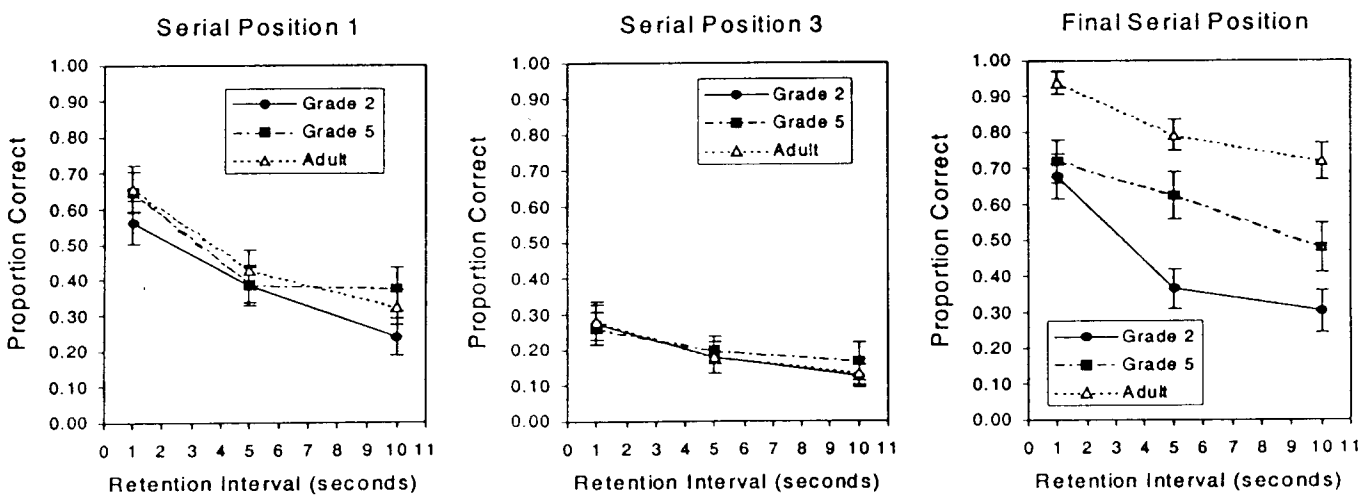


Fig. 8. Loss of memory for unattended digits within a list across retention intervals, selected here for illustrative purposes for Serial Position 1 (left-hand panel), Serial Position 3 (central panel), and the final serial position (right-hand panel) from the data set of Cowan, Nugent *et al.* (2000). Within each panel, age group is the graph parameter. Age differences in decay rates were obtained only at the final serial position.

this serial position turns out to be independent of the list length, the result perforce will be difficulty-insensitive. If not, it may be necessary to formulate a more complex function to isolate the part of final-serial-position performance based on a sensory memory that is independent of list length, such as the proportion decrement in performance across retention intervals (see above).

III. Conclusions, Observations, and Speculations

Working memory is a complex system of processes that retain information temporarily in the service of the performance of various cognitive tasks. Thus, Hulme and Roodenrys (1995, p. 374) aptly remarked that, in cognitive tasks such as language comprehension and mental arithmetic, "... it is obvious that a wide array of memory processes will be operating. Working memory is a convenient shorthand to refer to these processes and their functional importance."

We have suggested that the development of this complex performance can be understood better by breaking it down into basic processes that can be studied in detail. One difficulty that complicates such an approach is that individuals' basic processes can be identified not from a single data point, but only from a *pattern* of results across several conditions. One typically is faced with an unfortunate choice between comparing children's patterns of performance across ages on comparable stimuli, but at different levels of performance, or on stimuli that are not comparable because they have been individually adjusted to match performance levels. Neither option may yield a fully interpretable developmental comparison. We have suggested, though, that difficulty-insensitive measures might be derived, using theoretical considerations, for at least three basic types of parameters of working memory: capacity limits, retrieval rates, and decay rates. For these measures, which presumably reflect the underlying processes of interest, estimates of developmental change can be obtained that are stable across a range of difficulty levels of the stimuli.

Nevertheless, a number of important questions remain unanswered. We end by examining these questions and giving tentative answers to them.

A. ARE THERE TRULY DIFFICULTY-INSENSITIVE MEASURES OF WORKING-MEMORY PARAMETERS?

We have manipulated difficulty level in a very specific way, by altering the lengths of lists or the set sizes of arrays of to-be-remembered stimuli. Difficulty could be manipulated in other ways. For example, word frequency and word length both influence the level of recall and one could ask whether the three parameters we have examined are constant across levels of difficulty defined by such manipulations as opposed to manipulations of list length.

Answering this question is not easy. One first must ask whether any particular manipulation of difficulty preserves the conditions under which the parameter estimate theoretically can be derived. For example, the capacity parameter can be derived only if each item is an independent, single chunk of information. Recall of nonwords is poorer than the recall of words but that difficulty manipulation does not preserve the appropriate conditions, inasmuch as each nonword (or, say, a word with which the participant is unfamiliar) may be represented in memory as a collection of subitem parts rather than as an integrated, lexical item or single chunk.

Another possible scenario is that a parameter estimate could change with a change in the difficulty of the stimulus set because that parameter underlies the change in difficulty. Once more a word/nonword distinction is relevant. Hulme *et al.* (1999) found that interword pauses are shorter for lists of words than for lists of nonwords. One possibility is that this difference occurs because nonwords are not single, integrated units in memory (as mentioned above), in which case the effective list length is higher than the nominal list length and is unknown. Even if each nonword were, in some sense, a recently stored single unit in memory, the time needed to make contact with such nonword units in memory might be longer than the time needed to make contact with real words because of insufficient long-term memory support for the nonwords. This would be expected to affect memory-search time, altering the durations of interword pauses; and it would be possible that this difference in memory-search time could cause a difference in memory span. As another example, if stimulus items differed in their consonants but not their vowels in an unattended-speech procedure, this should increase the amount of memory decay relative to a situation in which vowel variation was included (Cowan, Lichy, & Grove, 1990). The poorer performance with no vowel variation could indeed be attributed to the decay of sensory information in the sense that less redundant, more fragile sensory information is available to work with in the first place.

What is one to conclude given such examples? Essentially, if a manipulation (such as list length) produces difficulty-insensitive measures, these measures can be used to make developmental comparisons. However, many factors can cause a manipulation to fail to produce difficulty-insensitive measures, and the existence of such cases does not invalidate the logic of difficulty-insensitive measures in general.

B. WHAT ARE THE RELATIONS BETWEEN THE STUDIED WORKING-MEMORY PARAMETERS?

One simplistic view would be that the capacity, retrieval, and decay parameters that we have observed are three independent influences on recall. We would caution that they may not be independent from one another. Perhaps not all of the parameters are causally related to memory; one might be causal and another

a noncausal correlate of the first. We have not yet examined this question directly but certain evidence may have a bearing upon it.

Cowan (1998) suggested that capacity contributes to memory span, whereas retrieval speed may not and may instead be another consequence of capacity. The finding underlying that proposal was that the mean interword pauses in serial recall were the same in 4- and 8-year-olds when examined at each age for span-length lists. However, the same pattern does not necessarily hold up in older children. Figures 6 and 7 show that pauses within responses for lists of length (span = 2), or 2 below maximal span, seem to speed up across age groups from third to fifth grades. From the same data set (Cowan *et al.*, 1998, Experiment 1), we found that pauses within responses for lists of length (span = 1) were longer for 22 children with a maximal span of 5 ($M = 341$ ms, $SD = 186$) than for 34 children with a maximal span of 6 ($M = 279$ ms, $SD = 179$). (More complete data, for longer lists within this experiment, were not estimated, partly because there are few trials for maximal span-length lists.) Thus, adjusting list length for ability does not eliminate the speed differences in recall. Another relevant consideration is that Cowan (1999b) found that fifth-grade (10- to 11-year-old) children had shorter interword pauses than third-grade children, even among a subgroup with spans matched across ages. These results suggest that interword pauses and the retrieval rates that they estimate are, at best, indirectly related to memory span. If capacity were the cause of memory span changes, we would predict that interword pauses would not be directly related to capacity measures either. The predominant role of capacity is reasonable inasmuch as Cowan, Nugent *et al.* (1999) found a fairly strong correlation between memory for attended digits and memory for ignored digits, $r = .52$. We have not yet examined direct correlations between difficulty-insensitive measures of retrieval rates and difficulty-insensitive measures of memory capacity, so we cannot definitively say whether they are independent.

Although we are proposing that retrieval rates do not directly influence spans, they may have an indirect influence. For example, sufficiently rapid retrieval rates may permit the use of a more mature strategy such as grouping of items together in memory, though rapid retrieval would not guarantee that the mature strategy would be used (Cowan, 1999b).

Capacity, retrieval rate, and decay rate all seem to change with age. For this reason, theoretically speaking, they must be correlated with one another; but it is unclear if the correlations would persist with age partialled out. Even if the measures were correlated within an age, more work would be needed to determine the paths of causality.

C. WHAT ARE THE BASES OF DEVELOPMENTAL CHANGE IN THESE WORKING-MEMORY PARAMETERS?

One could make a case for either biological or environmental factors leading to age changes in the parameters we have examined. Regarding biology, it has been

proposed that frontal-lobe structures underlie working memory (for reviews see Engle, Kane, & Tuohiski, 1999a; Nelson, 1995). However, the type of mechanism that may be reliant on the frontal lobe is the control of attention. Less certain are the biological structures that underlie the parameters of working memory that we have examined. Cowan (1995) provided rationales for the proposal that the focus of attention (and hence, by implication, its capacity) depends heavily on the inferior parietal areas, whereas decay of sensory information probably depends on the sensory cortical areas. Retrieval speed may depend on language-production-related parts of the frontal lobe.

Regarding environmental factors, one could argue that the greater familiarity with stimulus materials accumulated across ages in childhood underlies age differences in measured parameters. Less familiar materials could result in a lower capacity, longer retrieval times, and faster decay than more familiar materials. We have only one piece of evidence contradicting that interpretation at present. Cowan, Nugent *et al.* (1999a) noted that children receive much more exposure to the digits 1-3 than to the digits 7-9, yet there were no differences at any age in the recall of digits 1-3 versus 7-9. Perhaps, when very familiar materials are used, there is a ceiling level beyond which familiarity plays no role. Given that the target items in the studies discussed above were digits (or, in one case, simple colors to be matched), familiarity probably played only a small role, if any, in these results. Thus, the present results are at least promising as potential behavioral indices of biological, neurological change in working memory systems.

D. HOW SUCCESSFUL ARE THESE WORKING-MEMORY PARAMETERS IN ACCOUNTING FOR HIGHER LEVEL COGNITION?

Daneman and Carpenter (1980) and Daneman and Merikle (1996) suggested that the complex span tasks, in which processing and storage requirements are imposed simultaneously, are much more successful than simple span in accounting for various measures of achievement and aptitude because they tax both processing and storage components of working memory. Complex span tasks are more successful, but the reasons are not yet clear. Engle *et al.* (1999a) suggested that the storage plus processing requirement taxes the ability to control attention, whereas Towse *et al.* (1998) argued that, at least in children, it is the longer duration for decay, not task difficulty, that is important in complex span tasks.

More complex span task results may correlate with achievement and aptitude measures well because they reflect a variety of simpler skills, each of which contributes a smaller but at least partly nonredundant correlation with achievement and aptitude tests. If a simpler, more basic parameter is found to correlate with complex tests, therefore, it is in principle more impressive.

We do not yet have tests of the relations between the difficulty-insensitive measures and tests of intellectual achievement or aptitude. However, Mukunda and Hall (1992) provided some important evidence that

