The Microanalysis of Memory Span and Its Development in Childhood

Nelson Cowan, J. Scott Saults, Lara D. Nugent, and Emily M. Elliott
University of Missouri, Columbia, USA

It is not clear from the literature why, as children develop, there are important increases in memory span, the number of just-presented items that the participant can repeat in the correct serial order. To understand this, some recent results on capacity limits and processing rates were re-analyzed. We first describe results using a conventional measure of performance in immediate memory tasks that is affected by the list length (proportion correct). Next we describe results using a less conventional measure (number correct) that is unaffected by list length under circumstances in which attention to the list during its presentation is curtailed. This measure can estimate an individual's limited-capacity storage ability. Last, we examine measures of spoken response timing that do and do not change with list length. We show that unconventional measures that do not change with list length, but do change with development, are especially useful in assessing basic changes in information processing parameters, including increases in memory capacity and processing speed.

Il existe une augmentation importante de l'empan au cours du développement de l'enfant, mais les raisons de cet accroissement demeurent incertaines. Afin de comprendre les causes de ce phénomène, de nouvelles analyses ont été effectuées sur des résultats récents concernant les limites de capacité et la vitesse de traitement. Nous décrivons tout d'abord les résultats en utilisant une mesure conventionnelle de la performance qui est affectée par la longueur de la liste (proportion correcte) dans des tâches de mémoire immédiate. Nous décrivons ensuite ces mêmes résultats en utilisant une mesure moins conventionnelle (nombre correct) qui n'est pas affectée par la longueur de la liste dans des conditions où l'attention est réduite. Finalement, nous examinons les mesures d'organisation temporelle des réponses orales qui changent ou non avec la longueur de la liste. Nous montrons que des mesures non conventionnelles qui ne changent pas avec la longueur de la liste, mais qui changent avec le développement, sont particulièrement utiles afin d'évaluer les changements se produisant au niveau du traitement de l'information.

In memory span tasks, lists of words or digits are presented with a serial recall period immediately following each list, and what is measured is how long the list can be before the participant makes errors in recalling the list. Performance on memory span tasks increases markedly with development in childhood, is related to performance on a number of more complex comprehension and problem-solving tasks (Dempster, 1985), and is even included in tests of intelligence. This makes sense because comprehension and problem solving require that propositions or premises be kept in the mind until additional information is presented and all of the information can be processed together. Memory span tasks provide one index of the ability to hold multiple items in the mind at one time (another practical example being addends within a math problem).

Although memory span tasks are procedurally simple, it is far from clear what mental processes contribute to performance on these tasks and how they operate together. To answer that question, what would appear to be needed is a detailed analysis or "microanalysis" of performance on memory span tasks. Siegler and Crowley (1991) proposed a microgenetic analysis of the time course of developmental growth in cognitive tasks, and what we are proposing is, in a complementary fashion, a microanalysis of the *processes* involved in span and its development. Baddeley, Thomson, and Buchanan (1975) initiated one such approach by determining that people could recall about as many items as they could pronounce in about 2sec, and by suggesting a phonological loop process consisting of a fading, 2-sec passive phonological store along with an active rehearsal process that is used in an attempt to refresh information in the store before it fades away. Hulme, Maughan, and Brown (1991) demonstrated that there is an additional, independent role of lexical knowledge, with words being recalled better than nonwords that could be spoken just as quickly as the words. We fully acknowledge such factors but have found that there are still other factors that appear important and yet have been studied far less often with reference to memory span. We have studied (1) the role of attention in an immediate, serial verbal recall task

Requests for reprints should be addressed to Nelson Cowan, Department of Psychology, University of Missouri, 210 McAlester Hall, Columbia, MO 65211, USA (Tel: 573–882–4232; Fax: 573–882–7710; E-mail: CowanN@missouri.edu).
This research was supported by NIH Grant HD-21338. We thank Nathaniel Fristoe and an anonymous reviewer for comments.

(Cowan, Nugent, Elliott, Ponomarev, & Saults, in press), and (2) the speed of spoken recall in a memory span task (Cowan et al., 1998), and it is these experimental situations that we will re-examine here.

The main thesis of this article is that one can identify two different types of measures of immediate memory that change with development: those that are affected by list length and those that are unaffected by list length. The list length manipulation is a well-known way to influence the difficulty of a memory test, both in adult research (e.g. Raaijmakers & Shiffrin, 1981) and developmentally (e.g. Engle & Marshall, 1983). Moreover, for measures affected by list length, one can individually adjust the list lengths so as to equate the task difficulty level and reduce or eliminate age differences in performance. However, there are other, less conventional memory measures that are unaffected by list length and do not tap difficulty level per se, yet still reveal age differences. We show that such measures can index basic aspects of the processing capacities of individuals that are independent of the specific task parameters. One such aspect is the immediate memory storage capacity, which will be explored with respect to the memory for unattended speech procedure of Cowan et al. (in press). Another such aspect is processing speed, which will be examined with the spoken recall timing procedure of Cowan et al. (1998).

MEMORY FOR ATTENDED AND UNATTENDED DIGIT LISTS

The experiment by Cowan et al. (in press) showed that proportion correct and number correct measures yield different types of information about recall. First, let us describe the rationale and results of the experiment. Cowan et al. highlighted two components of memory span. One component is a capacity to hold in mind a certain number of unrelated items (e.g. three or four random digits). Another component is the use of attention to engage in mnemonic strategies (e.g. grouping and semantic elaboration) that enable the participant to chunk or link items together so that they are no longer unrelated (see Miller, 1956). Cowan et al. set out to separate these components in a participant sample including first-graders, fourth-graders, and adults (N = 24 per age group). In one part of the experiment, lists of spoken digits were attended, as in an ordinary span task. In another part of the experiment, lists of spoken digits were ignored as the participant carried out a silent video game in which a central picture on each trial was to be matched to one of four surrounding pictures on the basis of a rhyme between their names. Although there were no responses to most of the spoken lists, occasionally the video game was replaced on the computer screen with a series of empty boxes, one box per item in the most recent list, and the participant was to use the keypad to recall the digits from this most recent list in the presented serial order. The cue to recall came only 2sec after the most recent list. The lengths of the lists within the attended-speech and ignored-speech sessions varied and were adjusted to equal the maximal number of items that the child could repeat (termed Span Length), or one, two, or three items less (Lengths Span-1, Span-2, and Span-3).

It was expected that the response to unattended lists followed by a recall cue would have to be based on a vivid sensory memory of the list, as in the previous studies using visual spatial arrays (Sperling, 1960) and auditory spatiotemporal arrays (Darwin, Turvey, & Crowder, 1972). From these studies it can be assumed that the amount of information in auditory sensory memory is very large and extends backward several seconds in time. Therefore, memory for unattended speech should be limited not by this store, but by how much information could be transferred to a more categorical, limited-capacity store (e.g. the focus of attention; see Cowan, 1988) without exceeding its capacity. The results suggested that this analysis is appropriate. Sperling (1960) found that people could recall about 4 items in his "whole report" condition regardless of the array size and, similarly, Cowan et al. (in press) found that adults could recall about 3.5 items in the unattended speech condition no matter whether the list length was equal to Span, Span-1, Span-2, or Span-3. Children also displayed a constant memory for unattended speech across list lengths, but the limit was significantly smaller (with first-graders recalling only about 2.5 items and fourth-graders recalling about 3.0 items).

It is important to consider why it is impossible for participants to increase the number correct through mnemonic strategies such as covert verbal rehearsal or memorization as the display size increases. In Sperling's (1960) study, the whole report limit could not benefit from these strategies because the array sizes were too large, and the arrays presented too briefly, for participants to get a chance to use strategies. In our study the items were presented one at a time but, in the unattended speech condition, the absence of attention severely limited the use of strategies.

In contrast to unattended speech, the number correct for attended speech increased steadily as a function of list length. However, the age differences were about the same size as in the unattended speech condition, showing that age differences in recall have little to do with the use of attention-demanding mnemonic strategies during the presentation of the list. This finding is in contrast to the classic emphasis on the importance of rehearsal during reception of the list in short-term memory tasks (Flavell, Beach, & Chinsky, 1966; Ornstein & Naus, 1978), though it does agree with several other previous analyses of memory span that questioned the central importance of rehearsal in span tasks (Dempster, 1981; Huttenlocher & Burke, 1976).

We now consider what proportion correct and number correct scores tell us about the processes involved in our task. Figure 1 plots the results for the attended- and unattended-speech conditions in terms of the proportion correct. (An item was counted correct only if it was

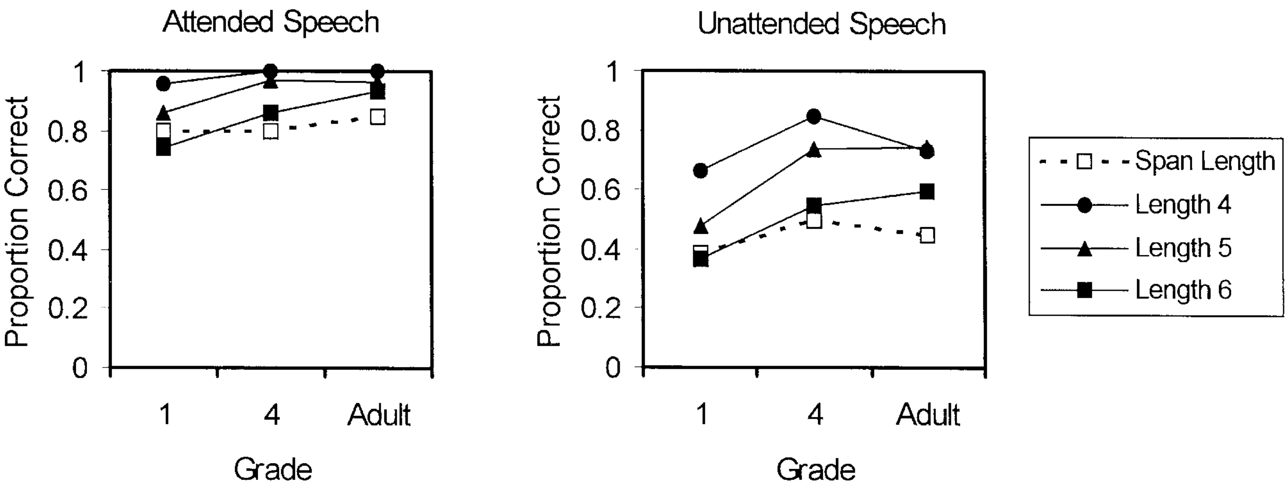

FIG. 1. For participants of three ages, proportion correct in two immediate memory tasks for three list lengths (solid lines) and for lists with a length equal to the maximum that the participant recalled correctly (dashed lines). *Left panel*: attended-speech task. [In that task, standard errors in the three age groups (youngest to oldest) were: for List Length 4, .01, .00, & .00; for List Length 5, .03, .01, & .01; for List Length 6, .03, .02, & .02; and for span-length lists, .03, .04, & .02.] *Right panel*: unattended-speech task, in which lists were unattended during their presentation and recalled only occasionally, if a post-list visual cue was presented. [In that task, standard errors in the three age goups were: for List Length 4, .05, .04, & .04; for List Length 5, .05, .04, & .04; for List Length 6, .04, .04, & .04; and for span-length lists, .04, .05, & .03.] Data from Cowan et al. (in press).

recalled in the correct serial position.) The left panel shows the attended condition. Although performance levels are high, they strongly conform to the list length dependent pattern. The solid lines show performance for List Lengths 4, 5, and 6, the only list lengths available for a large enough number of participants in each age group. Performance levels are higher for the shorter list lengths, performance increases as a function of age, and the age effect is largest for the longest list length. Most importantly, as shown by the dashed lines, the increase in proportion correct across ages is basically eliminated when each participant is examined with lists of his or her own span length. The right panel of Fig. 1 shows much the same thing for the unattended speech condition, at a lower level of performance. Thus, the proportion correct measure is nicely sensitive to the list length in relation to the participant's span.

From the proportion correct scores, one could be led to suppose that the processes taking place in the attended- and unattended-speech conditions are fundamentally the same, with just an overall impairment in the unattended-speech condition. Evidence against this belief, in the form of number correct scores based on the same data as the proportion correct, appears in Fig. 2. If the recalled items have to be drawn from sensory memory into a capacity-limited store, with little chance of rehearsal or memorization during the presentation of the list, then the number reported should not exceed the capacity of the store no matter how many items are in the list. This is not the pattern shown for

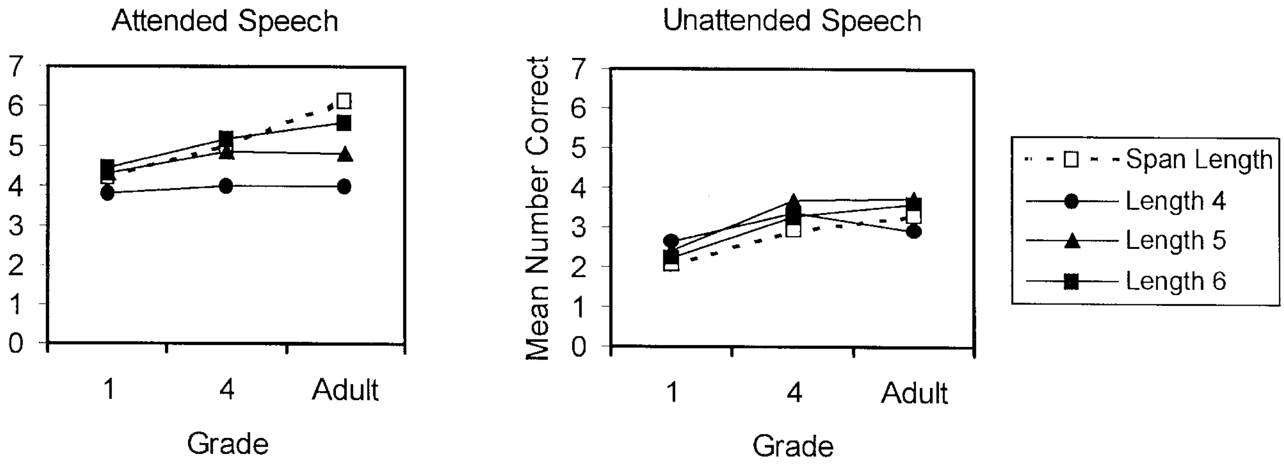

FIG. 2. For participants of three ages, number correct in two immediate memory tasks for three list lengths (solid lines) and for lists with a length equal to the maximum that the participant recalled correctly (dashed lines). *Left panel*: attended-speech task. [In that task, standard errors in the three age groups (youngest to oldest) were: for List Length 4, 0.05, 0.00, & 0.01; for List Length 5, 0.14, 0.05, & 0.06; for List Length 6, 0.30, 0.17, & 0.11; and for span-length lists, 0.17, 0.19 & 0.22.] *Right panel*: unattended-speech task. [In that task, standard errors in the three age groups were: for List Length 4, 0.19, 0.18, & 0.19; for List Length 5, 0.24, 0.23, & 0.23; for List Length 6, 0.42, 0.29, & 0.24; and for span-length lists, 0.19, 0.25, & 0.23.] Data from Cowan et al. (in press).

the attended-speech condition, depicted in the left panel of Fig. 2. There is still a clear list length dependence, and therefore no evidence that performance was restricted by a limited-capacity store. The list length dependence in this figure is as one would expect if some of the items were recalled from long-term episodic memory. Longer lists provide more opportunities to recall from long-term memory and therefore the number correct rises as a function of list length, with the largest list length effects in the most mature participants.

The number correct measure for the unattended speech condition, shown in the right panel of Fig. 2, tells a very different story. The number correct increases across ages, but that increase in the number correct is the same no matter which list length is used to examine it and, therefore, no matter whether the list length is adjusted to match the participant's span or not. We can think of the basic memory capacity that shows up in this condition as an age-dependent capacity that is rather distinct from what is observed in an ordinary span task. (The correlation between the pretested span and the mean number correct for unattended speech was r = .52, but the correlation between the pretested span and the number correct for attended speech was much larger at r = .94.) As noted, the number correct measure for unattended speech may reflect a limit to how many items can be drawn from sensory memory into the focus of attention at the same time. Assuming that far more than that number of items are available within an unanalyzed sensory memory, the limit in capacity of the focus of attention can be observed in the same way for all lists that contain at least as many items as the participant's capacity limit.

SILENT PERIODS WITHIN SPOKEN RESPONSES IN A DIGIT SPAN TASK

We now turn to a second area that can benefit from an empirical analysis similar to the one described earlier, in which a measure that is not affected by list length is derived. In an attempt to determine how the timing of spoken recall is related to recall itself in a digit span task, Cowan et al. (1998, Exp. 1) investigated several types of timing measures in first-, third-, and fifth-grade children (N = 24 per age). The article was focused on two types of measures: the duration of inter-word pauses within spoken responses in the memory span task, using only those trials in which the response was totally correct; and the duration of repetitions of short lists in a separate task in which the child was to speak as quickly as possible (speeded speech). These two types of measures proved to be very revealing because they both were correlated with memory span (at moderate levels of about .4), whereas they were not correlated with each other at all. Together, in a structural equation model including multiple measures of the two types of processing duration contributing to two separate latent variables, they accounted for 60% of the span variance and 87% of the age-related span variance.

Whereas the rapid speaking measures were assumed to reflect some articulation-specific ability such as verbal rehearsal capability (cf. Baddeley, 1986), inter-word pauses within the span task recall protocol were assumed to reflect the speed with which participants can search for information in the phonological memory trace to determine which item to recall next (Cowan, 1992). This conclusion is suggested by the pattern of responses across different list lengths and different word lengths. Each inter-word pause in the response is affected by the list length, as if the entire list enters into the processing in some way during these pauses (Cowan, 1992; Cowan et al., 1998). However, these pauses have been found to be unaffected by word length (Cowan et al., 1994; Hulme, Newton, Cowan, Stuart, & Brown, in press). The process accounting for our inter-word pauses thus did not appear to be rehearsal, which occurs at a slower rate for sets of words that take longer to pronounce (Baddeley et al., 1975; Baddeley, 1986; Landauer, 1962) and presumably would have produced a word length effect on pauses.

Other evidence strengthens the analogy between the processes taking place during inter-word pauses and during memory search tasks. Adults' rapid pronunciations of subspan lists (Sternberg, Monsell, Knoll, & Wright, 1978; Sternberg, Wright, Knoll, & Monsell, 1980) similarly have been explained according to a memory search process. An analogy with memory search is further strengthened in that, similar to the pauses in serial recall (but unlike serial recall performance levels), word length has no effect in standard memory search procedures in which a list is followed by a probe item (Chase, 1977; Clifton & Tash, 1973).

The work of Hulme et al. (in press), which used adult participants, further strengthens the hypothesis that inter-word pauses are times when some kind of memory search process takes place, in that memory search times (using a procedure in which a list was followed by a probe) were found to be correlated with inter-word pauses in the span task recall period. Hulme et al. further suggested that during the pauses, in addition to carrying out a type of memory search, participants also must identify the lexical node in long-term memory corresponding to the phonological representation, given that pauses were much longer for nonwords than for English words.

The list length dependence of inter-word pauses alluded to here is an indication that these pauses are somehow related to task difficulty. This can be seen in a plot of inter-word pauses averaged across serial positions (Fig. 3, left panel). In the Cowan et al. (1998) data, pauses were available for different list lengths for different participants. For all participants, they were available for List Lengths 2, 3, and 4 (solid lines); and for all but one participant (who had a maximal span of 8 items), they were available for a relative list length of Span-2 (dashed line), which equalled a List Length of 2 (for Span = 4), 3 (for Span = 5), 4 (for Span = 6), or 5 (for Span = 7).

Longer list lengths were recorded but the meticulous timing analyses were not carried out on them. The results depicted in the left panel of Fig. 3 show the typical list length dependent pattern in which performance was better (i.e. in this case, contained shorter pauses) for shorter lists and in which the age effect was diminished by examining each participant on a list length linked to his or her own memory span (in this case, Span-2).

An important issue related to inter-word pauses and other measures of processing speed is that it is often unclear whether speed is the cause of capacity differences, whether it is the result of these differences, or whether both are affected by a third factor. Exquisite fits between speeds and span, with other variables controlled, have tended to lead to the suggestion that processing speed limits cause capacity limits (e.g. see Fry & Hale, 1996; Kail & Park, 1994; Salthouse, 1996). However, given the multiplicity of processing speeds, as demonstrated for example by Cowan et al. (1998), it is possible that some processing speed limits have a different causal path than others. We will use a new measure to gain insight into the nature of the inter-word pause measure.

Assume for the time being that participants must in some sense process the entire list in a search process taking place during each inter-word pause in the recall period (cf. Cowan, 1992; Hulme et al., in press). Further assume that the processing time for each item does not depend on the difficulty of the task, but depends only on the participant's intrinsic speed of processing that is not influenced by task load. If these assumptions are met, then the processing speed *per item* should not vary with list length, whereas it should vary among individuals.

This proposition is examined in the right panel of Fig. 3. This panel shows each mean pause time (averaged across serial positions) divided by the list length, providing an estimate of the mean pause time on a per-list-item basis. As the figure shows, this type of metric eliminates list length effects. This suggests the plausibility of the notion that an individual has a certain per-list-item pause time that stays the same no matter how many items are in the list.

CONCLUSION

It can be argued that measures of immediate memory that produce a list length effect are potentially ambiguous in developmental studies. They display age effects that differ depending on what list length is examined. Therefore, it is not clear how to map the complex age effects that are found onto underlying theoretical variables. In the present study, we have illustrated the use of measures that are derived in such a way that they are not influenced by list length, but nevertheless are influenced by age. We are hopeful these measures might be useful in identifying basic processes underlying immediate memory performance. A measure of the number correct based on data from Cowan et al. (in press) showed a complex pattern for attended lists, but an exquisitely simple pattern for unattended lists (Fig. 2, right). In that pattern, the number correct was the same regardless of the list length and was taken to indicate a developmental growth in the capacity limit of the store to which sensory memory of items from ignored lists must be transferred to allow recall. A newly derived speech recall timing measure, of the pause duration per item in the list, showed a pause-per-list-item time that was fixed across list lengths but sped up with development (Fig. 3, right). This measure was taken as a possible indication of developmental growth in an underlying speed of processing. Although this conclusion requires the acceptance of

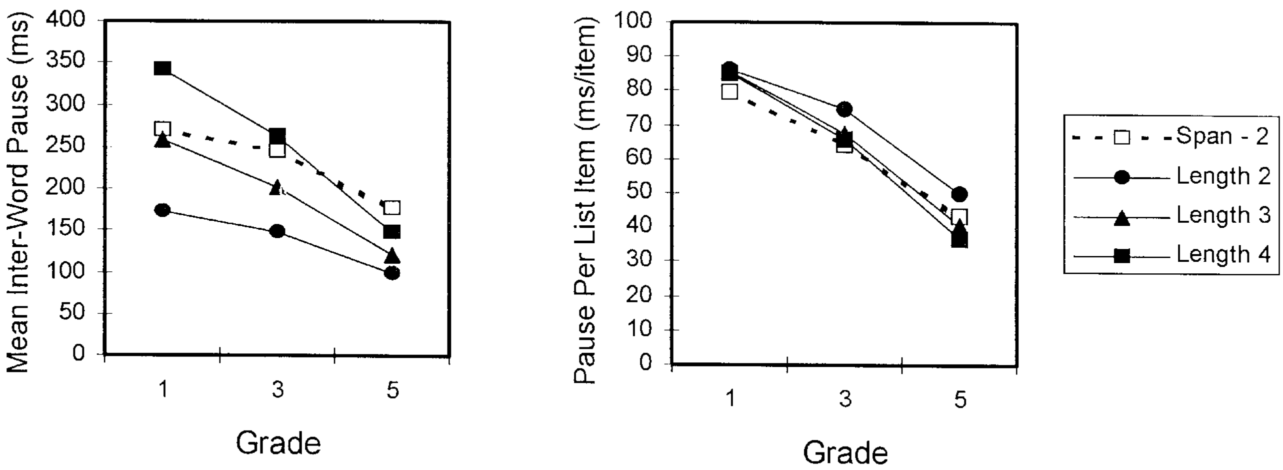

FIG. 3. For children of three ages, timing measurements within correctly repeated lists of three different fixed lengths (solid lines) and lists of a length two below the maximum that the participant recalled correctly (dashed lines). *Left panel*: inter-word pause times averaged across serial positions. [For this measure, standard errors in the three age groups (youngest to oldest) were: for List Length 2, 18.85, 23.98, & 20.95; for List Length 3, 26.93, 27.49, & 30.16; for List Length 4, 35.11, 28.77 & 32.75; and for List Length (Span-2), 27.11, 29.52, & 33.74.] *Right panel*: Pause per list item, calculated as the mean inter-word pause time (averaged across serial positions) divided by the list length. [For this measure, standard errors in the three age groups were: for List Length 2, 9.43, 11.99, & 10.48; for List Length 3, 8.98, 9.16, & 10.05; for List Length 4, 8.78, 7.19, & 8.19; and for List Length (Span-2), 8.09, 8.15, & 9.15.] Means calculated from the data set used by Cowan et al. (1998).

some uncertain assumptions about information processing that were described earlier, we feel that the present data and new measures warrant further investigation in the near future.

REFERENCES

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

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

Chase, W.G. (1977). Does memory scanning involve implicit speech? In S. Dornic (Ed.), *Attention and performance VI* (pp. 607–628). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Clifton, C., & Tash, J. (1973). Effect of syllabic word length on memory-search rate. *Journal of Experimental Psychology*, *99*, 231–235.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin*, *104*, 163–191.

Cowan, N. (1992). Verbal memory span and the timing of spoken recall. *Journal of Memory and Language*, *31*, 668–684.

Cowan, N. (1998). What is more explanatory, processing capacity or processing speed? [Commentary on target article by Graham Halford.] *Behavioral and Brain Sciences*, *21*, 835–836.

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

Cowan, N., Nugent, L.D., Elliott, E.M., Ponomarev, I., & Saults, J.S. (in press). The role of attention in the development of short-term memory: Age differences in the verbal span of apprehension. *Child Development*.

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

Darwin, C.J., Turvey, M.T., & Crowder, R.G. (1972). An auditory analogue of the Sperling partial report procedure: Evidence for brief auditory storage. *Cognitive Psychology*, *3*, 255–267.

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

Dempster, F.N. (1985). Short-term memory development in childhood and adolescence. In C.J. Brainerd & M. Pressley (Eds.), *Basic processes in memory development* (pp. 209–248). New York: Springer-Verlag.

Engle, R.W., & Marshall, K. (1983). Do developmental changes in digit span result from acquisition strategies? *Journal of Experimental Child Psychology*, *36*, 429–436.

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

Fry, A.F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science*, *7*, 237–241.

Hulme, C., Maughan,S., & Brown, G.D.A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, *30*, 685–701.

Hulme, C., Newton, P., Cowan, N., Stuart, G., & Brown, G. (in press). Think before you speak: Pause, memory search and trace redintegration processes in verbal memory span. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Huttenlocher, J., & Burke, D. (1976). Why does memory span increase with age? *Cognitive Psychology*, *8*, 1–31.

Kail, R., & Park, Y.-S. (1994). Processing time, articulation time, and memory span. *Journal of Experimental Child Psychology*, *57*, 281–291.

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

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

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

Raaijmakers, J.G.W., & Shiffrin, R.M. (1981). Search of associative memory. *Psychological Review*, *88*, 93–134.

Salthouse, T.A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*, 403–428.

Siegler, R.S., & Crowley, K. (1991). The microgenetic method: A direct means for studying cognitive development. *American Psychologist*, *46*, 606–620.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, *74* (Whole No. 498).

Sternberg, S., Monsell, S., Knoll, R.L., & Wright, C.E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G.E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 116–152). New York: Academic Press.

Sternberg, S., Wright, C.E., Knoll, R.L., & Monsell, S. (1980). Motor programs in rapid speech: Additional evidence. In R.A. Cole (Ed.), *Perception and production of fluent speech* (pp. 469–505). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.