In press 2010, *Developmental Psychology*

With Development, List Recall Includes More Chunks, Not Just Larger Ones

Nelson Cowan, Anna Hismjatullina, Angela M. AuBuchon, and J. Scott Saults

University of Missouri, Columbia, MO, USA

and

Neil Horton, and Kathy Leadbitter, and John Towse

Lancaster University, Lancaster, UK

Correspondence:

Nelson Cowan

Department of Psychological Sciences

University of Missouri

18 McAlester Hall

Columbia, MO 65211

Tel. 573-882-4232

Email CowanN@missouri.edu

Main text: 8,953 words, 3 tables, 6 figures

Running head: CAPACITY AND CHUNK FORMATION

Original submission 5 May, 2008; resubmissions 17 February, 2010; 8 May, 2010

Abstract

The nature of the childhood development of immediate recall has been difficult to determine. There could be a developmental increase in either the number of chunks held in working memory, or the use of grouping to make the most of a constant capacity. In three experiments with children in the early elementary school years and adults, we show that improvements in the immediate recall of word and picture lists come partly from increases in the number of chunks of items retained in memory. This finding was based on a distinction between *access* to a studied group of items, i.e., recall of at least one item from the group, and *completion* of the accessed group, i.e., the proportion of the items recalled from the group. Access rates increased with age, even with statistical controls for completion rates, implicating development of capacity in chunks.

With Development, List Recall Includes More Chunks, Not Just Larger Ones

A continuing, basic riddle regarding cognitive development is just why older children and adults carry out immediate memory tasks more successfully than younger children. This type of age difference, as in the repetition of lists of words or digits, has long been considered a key hallmark of the cognitive development of the child (Binet & Simon, 1916/1980; Bolton, 1892; Jacobs, 1887; Ramsay and Reynolds, 1995; Wechsler, 1944, 1991; Wundt, 1906). Immediate memory performance also predicts individual differences in cognitive aptitudes among young elementary-school children (Cowan et al., 2005). The viable hypotheses regarding what factors underlie the development of immediate memory include covert verbal rehearsal abilities (Cowan, Saults, & Morey, 2006; Flavell, Beach, & Chinsky, 1966; Hitch, Halliday, & Littler, 1989; Hulme & Tordoff, 1989; Ornstein & Naus, 1978), processing efficiency (Case, Kurland, & Goldberg, 1982; Chi, 1976; Hitch et al., 1989; Kail, 1997); world knowledge (Chi, 1978; Huttenlocher & Burke, 1976; Kail, 1997; Thorn, Gathercole, & Frankish, 2005) and, most relevant to the present article, two other factors: (1) basic working memory capacity measured in chunks or slots available for those chunks (Burtis, 1982; Case, 1995; Cowan, Nugent, Elliott, Ponomarev, & Saults, 1999; Cowan et al., 2005; Pascual-Leone, 1970, 2005) and (2) the ability to group items together to form larger chunks or to make use of such grouping (Dempster, 1978; Harris & Burke, 1972; Towse, Hitch, & Skeates, 1999).

We address the latter two factors, which arose from the seminal work of George Miller (1956) on capacity constraints in immediate memory. In reviewing studies of adults, Miller found that subjects could recall about seven items. An item or chunk, however, could be any unit that was meaningful to the individual, and this depended on that individual's long-term memory structure. For example, individuals with an ordinary command of English typically could remember about seven digits, letters, or words. If, however, an individual knows certain acronyms (such as FBI, IBM, IRS, for the Federal Bureau of Investigation, International Business Machines, Internal Revenue Service, and so on), that individual displays greatly increased memory for letters forming such items. Subsequent work has shown that, when covert rehearsal is prevented or the materials are too long to be rehearsed, immediate memory is closer to three or four rather than seven items (Cowan, 2001). The key point is that retrieval of information from long-term memory can result in multiple stimulus items (such as letters) being recoded to form a single mental item (such as a three-letter acronym).

Returning to the issue of which factors contribute to the development of immediate serial recall performance, and applying the insights from Miller (1956), there could be an increase in either the number of working memory slots or the amount of information captured by the chunk occupying each slot (see for example Pascual-Leone, 1970 for a model of the former type; Case et al., 1982 for the latter). It is difficult to discriminate between these alternatives because there is no ideal measure of how much information is entered into each slot. Some previous work has addressed this issue among children by including spatially or temporally separated groups of items in the list (Harris & Burke, 1972; Towse et al., 1999) or by introducing pairs of identical items that would be very natural to group together (Burtis, 1982). However, these articles differ in their conclusions. Towse et al. concluded that the ability to use temporal grouping information increases with age in childhood and, convergent with this, Cowan, Elliott et al. (2006) found that adults but not second grade children reported that they mentally grouped items together from a standard digit span sequence. In contrast, Harris and Burke (1972) found that grouping cues helped children as young as second grade. Burtis was able to get grouping using simple stimuli and concluded, on the basis of the number of groups recalled, that there is a childhood developmental increase in the number of slots available in working memory. Further study is necessary.

We believe that an important opportunity to study the development and deployment of capacity comes from recent adult research into the formation of temporary and arbitrary bindings through "associative training" or "chunk familiarization." Cowan, Chen, and Rouder (2004) manipulated the temporary strength of item associations by presenting word stimuli to adults either alone as singletons, or in consistent pairs; the training schedule varied in the number of paired presentations. Subsequently, list recall and cued recall were used to determine how pair familiarization affected memory capacity. The principal finding was that, although adults' recall of the trained pairs was indeed better than of singletons (i.e., more words were recalled from trained pairs), the number of recalled chunks (learned pairs or singletons) nevertheless did not appear to change with chunk size (see also Chen & Cowan, 2005, 2009; Marmurek & Johnson, 1978; Slak, 1970; Tulving & Patkau, 1962). This finding was used to support the 'constant capacity' hypothesis of Cowan (2001), whereby the chunk capacity of memory is said to be stable within an individual even though usage efficiency increases when there are multi-item associations. The present research examines the extent to which children and adults can benefit from familiarization with multi-item chunks. We use this information to clarify the development of capacity in chunks.

To measure participants' sensitivity to the structure of lists and the presence of chunks, we further capitalized on the distinction between concepts of chunk *access* and chunk *completion* that were devised by Naveh-Benjamin, Cowan, Kilb, & Chen (2007). Their work demonstrated that adult aging produces a large decrease in the number of slots in working memory, coupled with a more mild decrease in the ability to form and use associations between items to recall larger chunks. If at least one item from a familiarized pair was recalled, it was assumed that the subject had access to that pair. The second measure was how often the accessed pairs were completed; this refers to the proportion of items recalled from accessed pairs. Naveh-Benjamin et al. found that older adults had substantially and consistently lower pair access scores than young adults. There also were slightly lower pair completion scores in older adults, in some conditions. These findings suggest that the principal relevant effect of aging on working memory was on the number of slots, rather than on the efficiency with which those slots are filled with information. The results also suggested that young adults may also have a slightly better ability to form new pairs on line when singletons that have not previously been paired are presented within a word list.

The access and completion measures can also be calculated for sets of more than two items; access still refers to the recall of at least one substantive word from the set, and completion still refers to the proportion of items recalled from accessed sets. Gilchrist, Cowan, and Naveh-Benjamin (2009) found a particular type of chunk, the short, simple English sentence, yielding completion rates of about 80% in both children in the early elementary school years and adults. For lists of unrelated simple sentences, nevertheless, the children accessed fewer sentences than did the adults, suggesting that the children could retain fewer chunks.

The present experiments extend the measures of chunk access and chunk completion to the study of the childhood development of word and picture list recall, which is important given the common use of these kinds of recall as psychometric measures of aptitude and as indicators of working memory capability. Given the complexities of psycholinguistic research with sentences, moreover, it is important to confirm the finding of a childhood developmental increase in chunk capacity using simpler materials like word and picture lists.

Experiment 1

This experiment adapted the pair-training method of Naveh-Benjamin et al. (2007) to an investigation of child development. Like them, we used serial recall to minimize individual differences in recall strategies but are nevertheless interested only in item recall; the complex, unresolved issue of what determines the serial order of chunks in list recall (e.g., see Farrell & Lewandowsky, 2004) is beyond the scope of our hypotheses regarding developmental changes in capacity and chunking ability.

Method

Design

In the *familiarization* phase, participants heard and repeated 28 disyllabic nouns, half of them as singletons and the other half as pairs. The set was repeated 4 times, with the words in a different order but with the same word pairings each time. Next, in each trial of the *cued recall* phase, a word was presented (a singleton or the first word from a pair) and the participant had to indicate whether it was paired with another word and, if so, what that word was. This procedure was repeated with the entire set of words from the familiarization phase 4 times, with feedback on each trial indicating the correct response, unless the participant was 100% correct during an earlier test cycle, after which the cued recall phase ceased. (No participant was excluded due to performance in this phase of the experiment.) Finally, in each trial of the *list (sequence) memory* phase, the participant was to listen to and recall in the presented order a list of 6 or 8 words. These lists included one 6-word and one 8-word list in each of three conditions: lists composed of (1) familiarized pairs, (2) familiarized singletons, and (3) novel, unstudied words that had not been included in the familiarization or cued recall phases. With this arrangement, each word from the familiarization phase was used once in a list and, across all three list conditions, 42 words were used, with no word repeated more than once in the lists. Key procedural details are summarized in the left-hand column of Table 1.

Participants

Data in this experiment were collected in Columbia, Missouri, USA. Adults (n=30) included 14 men and 16 women, with a mean age of 39 years and 6 months (SD = 16 years; range 17 to 66 years), participating for course credit. Children (n=30) included 17 boys and 13 girls, with a mean age of 8 years and 2 months (SD = 5.6 months; range 7;4 - 9;2), who received payment and a book. Participants were native speakers of English with normal or corrected-to-normal vision and no known hearing loss. Socioeconomic status information was unavailable.

Stimuli and Apparatus

From the MRC Psycholinguistic Database (Fearnley, 1997), 42 disyllabic nouns suitable for our participant sample age were selected, with Thorndike-Lorge frequencies ranging from 21,430 (*smile*) to 60 (*penguin*). The words were digitally recorded using an adult male voice and all words lasted approximately 0.5 s. They were presented using E-Prime (Schneider, Eschman, & Zuccolotto, 2002) at a comfortable listening level, through a notebook computer with built-in high-quality stereo speakers. There were three different assignments of words to conditions and each participant received one of the three assignments.

- - - - Table 1 here - - - -

Procedure

Each participant was tested individually in a quiet room. An item familiarization phase was followed by cued recall and then list memory phases.

Item familiarization. Words were presented either alone or in pairs, in the latter case with consistent pairings maintained throughout the experiment. The presentations included all of the words needed for the subsequent list recall phase except the novel words. The order of all presentations was random with the restriction that no word occurred on immediately successive presentations. The inter-word silent interval within a pair was 500 ms. The presentation of each singleton or pair was followed by 500 ms of silence and then a subsequent beep as a cue for the participant to repeat aloud the presentation that was just heard. The participant had 2 s to repeat the presentation.

Familiarization began with a 2-trial practice using words not to be used later, to ensure that the instructions were clear. Then the 28 words to be used in subsequent phases of the study were presented and repeated. The familiarization phase took approximately 5 min. Within that time, each singleton and each pair was presented 4 times.

Cued recall. This involved presentation of a word that was either the first item in a learned pair or a singleton, with presentations randomized. For paired words, the participant was to respond by saying the second word if they recognized that the presentation came from a known pair. For words that had occurred as singletons in the familiarization phase, the participant was to respond “no answer.” After the experimenter entered the response into the computer, it responded with the spoken feedback "good" for correct responses and, for incorrect responses, it said "no" and then repeated the pair or, for singletons, said "no answer."

The computer tested each pair and each singleton from the familiarization phase once. If any errors were made, this entire stimulus set was tested again (including items that had been recalled correctly or incorrectly) in a different random order of pairs, but using the same pairings; this process repeated until a set was recalled without error or a maximum of 4 rounds of presentation of the set was reached.

List (sequence) memory. In this phase, each list was presented acoustically and immediate, spoken serial recall was required. Words within a list were drawn from just one of the three familiarization conditions (novel words, i.e., a no-familiarization control condition; familiarized singletons; or familiarized pairs). The six lists, which occurred in a random order, included two per condition: one 6- and one 8-word list.

Within a list to be recalled, the words were always temporally grouped into pairs, even in the singleton and novel-word conditions. The list was presented with 500 ms of silence between words in a pair and 1500 ms of silence between pairs. After the last word pair and a 1500-ms silent delay, there was a beep to prompt recall. If the participant did not remember a word, he or she was to say "blank" and skip to the next one. If the participant failed to say "blank" or otherwise indicate skipping a word, the words were entered into successive response positions.

Results

Cued Recall

On the first round of cued recall, which everyone completed, the proportion correct in adults was .87 (SD=.15) and in children was .71 (SD=.12). There were 9 of 30 adults and 17 of 30 children who achieved less than 100% correct on the final, fourth cycle of cued recall. They were still included in the results. Adults received a mean of 2.50 cycles (SD = 1.28) of the cued recall set, compared to children's mean of 3.63 cycles (SD = 0.72). Nevertheless, the knowledge of word pairings was well equated across age groups. Thus, in the last cycle of cued recall testing for each participant, the average proportion correct was .93 in both the adults (SD = .12) and children (SD = .08).

List Recall

Consistent with the aims discussed above, we used a free scoring in which items recalled from the list were counted as correct regardless of the locations of these words within the response protocol, and we then used these scores to calculate access and completion rates averaged across serial positions.[1] Access refers to the mean proportion of pairs from which at least one item was recalled, and completion refers to the mean proportion of items recalled from accessed pairs. Thus, they provide separate sources of information about recall. We will use these measures first to examine the benefits of familiarization for list recall in the two age groups, and then to compare the chunk capacity in the two age groups.

Benefit of set familiarization. In order to examine the benefit of item familiarity and pair familiarity in each age group for our measures, we defined a *set* of items in this experiment as a pair of two adjacent items presented in close proximity in the list to be recalled; Items 1 and 2 formed the first set, Items 3 and 4 formed the second set, and so on. Familiarized pairs always corresponded to the pairs presented in list recall. By defining sets according to the list stimuli, we could find out whether familiarization with singletons or pairs affected access to and completion of the sets.

For an analysis of set access, the within-participant factors included list length (6 or 8 items) and familiarization condition of the items included in the list (unstudied novel singletons, familiarized singletons, and familiarized pairs). Overall, adults accessed 0.83 (SEM = .02) of the sets presented within lists whereas children accessed considerably fewer, 0.69 (SEM = .02) of the sets, a significant difference, $F(1,58)=28.11$, $MSE=0.07$, $\eta_p^2=0.33$.

There were two other main effects in the analysis, which can be seen in Figure 1 and, for the paired conditions, in Table 2. There was an effect of list length, $F(1,58)=65.09$, $MSE=0.03$, $\eta_p^2=0.53$, and an effect of familiarization condition, $F(2,116)=9.36$, $MSE=0.05$, $\eta_p^2=0.14$. The proportion of sets accessed was higher for 6-item lists than for 8-item lists. It was highest for the familiar pairs, lower for the familiar singletons, and lowest for the novel singletons. Newman-Keuls tests showed that pair access in all of the familiarization conditions differed from one another.

- - - - Figure 1 and Table 2 about here - - - -

The effect of age group did not interact with list length in the access scores. However, there was a marginal interaction of the age group with the familiarization condition, $F(2,116)=3.04$, $MSE=0.05$, $\eta_p^2=0.05$, $p=.052$. Newman-Keuls tests showed that none of the familiarization conditions differed from one another in adults, whereas all of the familiarization conditions differed from one another in the children.

Completion refers to the proportion of items recalled from accessed sets. We included the same factors as in the analysis of access. Five of the children had to be excluded from this analysis because they did not access any pairs in at least one condition of the experiment (i.e., recalled no items), making the completion score impossible to calculate. Overall, adults completed 0.85 of the accessed sets (SEM=0.01), whereas children completed only 0.80 of them (SEM=0.01), a significant difference, $F(1,53)=12.83$, $MSE=0.02$, $\eta_p^2=0.19$. Notably, age group did not interact with familiarization condition ($F<1$), so the benefit of studying items or pairs before the list recall test was no greater for adults than for children.

As shown in Figure 2 and, for the paired conditions, in Table 2, there were significant effects of list length, $F(1,53)=10.44$, $MSE=0.01$, $\eta_p^2=0.16$ and familiarization condition, $F(2,106)=39.70$, $MSE=0.02$, $\eta_p^2=0.43$. Set completion was somewhat higher for 6-item lists than for 8-item lists. It was highest for familiar pairs and much lower for familiar singletons and novel singletons. Newman-Keuls tests showed that completion for the familiar pairs exceeded the other two conditions, which did not differ, demonstrating the powerful effect of paired associate learning on the formation of chunks that were then used in list recall.

- - - - Figure 2 about here - - - -

The list length interacted with the familiarization condition, $F(2,106)=7.24$, $MSE=0.02$, $\eta_p^2=0.12$. Figure 2 shows that for the 8-item lists, familiarization with singletons actually reduced the amount of completion relative to unfamiliar singletons. This pattern is unlikely to reflect chance fluctuation inasmuch as the pattern was similar in both age groups, with no interaction with age group. Learning that an item is a singleton may make it more difficult subsequently to associate such an item with other items in the list.

Last, we also examined the relation between cued recall cycles and the ability to recall lists of paired associates. The partial correlation between the number of cued recall cycles and completion of accessed sets in the paired-associate condition (with age group partialled out) was significant for 8-item lists, $r=-.33$, and it was marginal for 6-item lists, $r=-.22$, $p=.10$. These correlations are such that individuals with more cued recall cycles had lower completion rates. Importantly, the corresponding correlations for access rates did not approach significance, $p>.1$.

Chunk capacity. This section will treat only the lists comprising familiarized pairs. Without further statistical control, estimates of pair access do not yield a clear indication of chunk capacity. This point can be illustrated as follows. Suppose that one individual recalls Items 1, 2, and 3 in a list whereas another individual recalls Items 1, 2, 3, and 5. It is possible that both individuals nevertheless have used same amount of capacity (say, 3 chunks) but that only the second individual was capable of recalling Items 1 and 2 as a single, learned chunk, allowing that individual to free up one chunk in working memory to retain another item in the list. Accordingly, in order to compare the capacities of two individuals, access scores must be examined with a statistical control for completion scores. For this purpose we used analysis of covariance (ANCOVA) on pair access by age group, using list length as a within-participant variable and using pair completion at the two list lengths as covariates. This approach cannot indicate the magnitude of an effect of age group on pair access, but it can indicate that there is such an effect. Given that the covariates are entered into the regression equation before age group, any variance shared between age group and completion scores will be credited to completion scores rather than to age group (Miller & Chapman, 2001, p. 45). This analysis therefore provides a conservative estimate of age group effects on access scores. This analysis produced no effects of the covariates, but there was an effect of age group, $F(1,56)=4.63$, $MSE=0.03$, $\eta_p^2=0.08$. The same was true when we also added the maximum performance from any one cycle of cued recall, $F(1,55)=4.80$, $MSE=0.03$, $\eta_p^2=0.08$.

Discussion

This experiment resembled the word-pair experiment that Naveh-Benjamin et al. (2007) used to examine aging effects in list recall. Naveh-Benjamin et al. found that young adults had a substantially higher capacity than older adults, as estimated by access to pairs within lists. Young adults also had a slightly better ability to form new familiar associations between items in the list, as estimated by the rates of completion of pairs that were accessed for the singleton condition. Our findings with childhood development were comparable. Eight-year-olds demonstrated poorer access to pairs than adults (Figure 1) as well as poorer completion of pairs than did adults (Figure 2). Nevertheless, the benefit of pair familiarization in the present study was about as great for the children as for the adults. This, too, is comparable to what Naveh-Benjamin et al. found for adult aging.

The differences between age groups actually appear larger for the novel and familiar singleton conditions than for the familiar pair conditions (Figures 1 and 2). This suggests that adults have some way to handle lists of unstructured items better than children, such as rehearsal and on-line grouping (e.g., Cowan et al., 2006). What is of greater relevance to the present research goal is to understand the residual differences between age groups in access for lists with familiarized word pairs. The ANCOVA approach provided strong evidence for a difference in chunk capacity. In the familiarized-pair conditions, adults accessed more pairs than children even with pair completion statistically controlled.

Experiment 2

In the remaining two studies, a picture-recall procedure is used with 6-item lists (Experiment 2) and 9-item lists (Experiment 3) in order to investigate further the developmental difference in chunk capacity. The subsequent experiments differed from the first in a number of ways so that convergence between the results of the experiments would indicate some generality in the phenomena. Pictorial instead of spoken stimuli were used (with learned labels for the pictures) inasmuch as the ability to form groups voluntarily appears to be stronger for visual stimuli than for sounds (Cowan, Saults, Elliott, & Moreno, 2002; Frankish, 1985). The list length was restricted to 6 items. To provide further validation of the concept that learned associations result in chunks within the list, we included a condition in which triplets, and not only pairs, of items were the sets of items associated before list recall. The general expectation is that the proportion of sets accessed will be higher for triplets (because there are fewer of them) but that the completion of accessed sets will be higher for pairs (because each set sets contains fewer items). Thus, with list length held constant, in the pair and triplet conditions, we can make the test more challenging for chunk access versus chunk completion, respectively.

There were other differences between the first two experiments. The cued-recall phase and the novel word lists were eliminated in this experiment, in order to leave time for a lengthy study period that included training on item labels and then familiarization with item singletons, pairs, and triplets. Further leaving the grouping structure to the active processing of the participant, the 6 list items were always presented in a regular pacing, not in pairs as in Experiment 1.

Method

Key procedural details are summarized in the right-hand column of Table 1.

Participants

Adults (n=21; 18 women and 3 men) with a mean age of 20 years, 11 months (SD = 3 years; range 18 - 29 years) were students at Lancaster University, U.K., who were paid £2 for their participation. All were competent English speakers. Children (n=21; 16 girls and 5 boys) with a mean age of 10 years, 1 month (SD = 8.2 months; range 9;0 to 11;2) volunteered to take part after parental consent had been obtained and were tested in their school environment. They included 16 girls and 5 boys, with a mean age of 10 years, 1 month (SD = 8.2 months; range 9;0 to 11;2). All were native English speakers. All participants had normal or corrected-to-normal vision and no known hearing loss. Socioeconomic status information was unavailable.

Stimuli and Apparatus

Fifty-four colored pictures of common nouns (animate and inanimate objects) suitable for children were drawn from a normed set taking into account both word frequency and picture identifiableness (Rossion & Pourtois, 2004). Objects appeared within a 112 x 84 mm screen window. All events were controlled by a Psyscript experimental control environment (Version 2.1.1, 2007, Simon Slavin, http://www.psych.lancs.ac.uk/software) under the Apple OS X operating system, and oral responses were recorded digitally.

Procedure

The experiment was divided up into three trial blocks. Each trial block involved object picture familiarization followed by list memory testing using those same picture configurations.

Object picture familiarization. Each trial block began with the presentation of 18 pictures organized into 11 presentations: six singletons, three pairs, and two triplets. The 11 presentations were then repeated a second time, in a different random order but with the same multi-item chunk configurations. The participant always responded to each presentation by naming the picture or pictures. Singletons were displayed center-screen for 2 s; paired pictures were displayed to the left and right of center for 4 s, and triplets were displayed in left, center, and right positions for 6 s. Thus, the rate of presentation was 2 s per picture, plus an interval of approximately 100 ms between displays as the software retrieved the object file. The assignment of stimuli to chunk size condition (singletons, pairs and triplets) in this phase was randomized for each participant, as was the order in which the objects were displayed.

List (sequence) memory. Participants were next told that six-object sequences would be presented using images from the preceding phase. Participants were now not to name the objects as they appeared but to try and recall them in the correct order as soon as the recall signal appeared. In each trial block, the experimenter initiated each trial, always comprising a sequence of 6 consecutive, isolated objects presented for 2 s each (with an inter-item blank interval of 100 ms) followed by the words "Recall now" displayed center-screen. Participants recalled the sequence verbally. One trial per block involved the 6 familiarized singletons, one involved the items from 3 familiarized pairs, and one involved the items from the 2 familiarized triplets. Trial order was randomized. Items that had been familiarized in multi-item sets were presented within the list one at a time, but in the same order in which they had appeared in the chunk during familiarization (assuming that left-to-right order in familiarization translates into first-to-last order in the memory trial). The order of pairs and of triplets in a list was unrelated to their order in the familiarization phase.

Results

Benefit of set familiarization. Analyses of sets accessed were carried out separately to highlight familiarized pairs versus familiarized triplets. In the pair condition, the sets were defined as Items 1-2, Items 3-4, and Items 5-6; in the triplet condition, the sets were defined as Items 1-3 and Items 4-6. In each analysis, to test the benefit of familiarization, lists of familiarized singletons were used as a control condition. They were analyzed as if adjacent items had been familiarized as two- or three-item sets (which they had not). Lower completion rates for singletons than for the familiarized pairs or triplets with which they were compared would indicate a benefit of set familiarization.

The results of analyses of the access measure are shown in Figure 3. In each analysis, a within-participant factor was whether or not the sets had been familiarized as such, or only familiarized as singletons. In the analysis of lists for pairs, the proportion of sets accessed was higher for adults than for children, $F(1,40)=24.72$, $MSE=0.01$, $\eta_p^2=.38$, but there was no effect of familiarization and no interaction of familiarization and group. Overall, adults accessed .98 of the sets (SEM=0.02), whereas children accessed only .86 of them (SEM=0.02). Similarly, in the triplet condition, the proportion of sets accessed was higher for adults than children, $F(1,40)=7.81$, $MSE=0.01$, $\eta_p^2=.16$, but the other effects were not significant. Overall, adults accessed .99 of the sets (SEM=0.01), whereas children accessed fewer, .95 of them (SEM=0.01).

- - - - Figure 3 about here - - - -

Recall was next scored according to what proportion of items was recalled from accessed sets, i.e., set completion. The results are shown in Figure 4. Separate analyses were again carried out for familiarized pairs and their singleton controls, and for familiarized triplets and their singleton controls. In both cases, there was a large main effect of age group with more set completion by adults than by children. For pairs, $F(1,40)=20.23$, $MSE=0.01$, $\eta_p^2=0.34$ and for triplets, $F(1,40)=28.67$, $MSE=0.01$, $\eta_p^2=0.42$.

- - - - Figure 4 about here - - - -

In both cases, unlike the access scores, there was a large main effect of the grouping condition for completion scores, with grouped (pair or triplet) familiarization resulting in more completion than singleton familiarization for both pairs and triplets (see Figure 4). For pairs, $F(1,40)=21.49$, $MSE=0.01$, $\eta_p^2=0.35$; for triplets, $F(1,40)=17.03$, $MSE=0.01$, $\eta_p^2=0.30$. Resembling Experiment 1, the familiarization condition did not interact with age group ($F<1$ for both pairs and triplets), indicating that the benefit of familiarization was comparable across age groups.

Chunk capacity. As in Experiment 1, capacity was examined in ANCOVAs designed to examine effects of access to familiarized sets with the completion rates controlled. For familiarized pairs, there was a significant age effect, $F(1,39)=5.34$, $MSE=0.01$, $\eta_p^2=.12$, but no effect of the completion scores used as the covariate. For triplets, neither the effect of age nor the effect of the covariate approached significance but this is may be due to a ceiling effect.

Discussion

The present experiment involved 6-item lists and again we found that the benefit of set familiarization for completion rates was comparable across age groups, as was also found for linguistic materials by Gilchrist et al. (2009). Gilchrist et al. actually found the same chunk completion rates across age groups but here, in contrast, we found lower completion rates in children. Nevertheless, we found equal benefits in children and adults of set familiarization relative to the singleton conditions.

Despite very high set access levels, an age effect in the access to pairs was obtained when pair completion was used as a covariate; this we take as further evidence of the capacity difference observed also in Experiment 1. For triplets, this was not the case, an unsurprising result given that there were only 2 sets to be accessed.

Experiment 3

The third experiment was designed to overcome two limitations of the previous experiments. First, the six-item lists used in that experiment were not long enough to examine access to triplets without encountering ceiling effects. In the third experiment, therefore, we modified the methods of Experiment 2 for an examination of the recall of nine-item lists, consisting of 9 familiar singletons or 3 familiar triplets.

Second, although our interest is in age differences in a core working memory capacity, we have not ruled out the contribution of rehearsal, given previous observations that the ability to carry out covert verbal rehearsal to assist recall improves markedly during childhood development (e.g., Flavell et al., 1966; Ornstein & Naus, 1978). To examine the contribution of rehearsal, we included a second group of adult participants who carried out a standard rehearsal-suppression technique (e.g., Baddeley, 1986) during the list memory task. It is likely that rehearsal suppression will impair performance compared to the no-suppression adults. The critical question would then be whether that impairment affects triplet access, triplet completion, or both factors; and whether it affects chunk capacity, as indexed by access with statistical control of completion.

Method

This sample was collected in Columbia, Missouri, USA. Incentives for participation were as in Experiment 1. The sample of 24 participants per group included children with a mean age of 8 years and 3 months, SD=4 months (10 girls, 14 boys), college students who carried out the experiment in a manner like the children, with a mean age of 19 years and 7 months, SD = 13 months (9 women, 15 men), and college students who carried out the list recall task with rehearsal suppressed, with a mean age of 18 years and 10 months, SD=5 months (15 female, 9 male). An additional 8 children were eliminated: 4 for experimenter error, 1 who became ill, 1 whose session was interrupted by his mother, and 2 who had untreated attention deficit disorder. Two adults were eliminated because of experimenter error. Socioeconomic status information was unavailable.

The method was the same as in Experiment 2 except for the number of pictures and the number and types of list memory trials, which were increased to allow more list recall trials. In this experiment, 72 pictures were used from the normed set (Rossion & Pourtois, 2004), 36 pictures in Trial Blocks 1-3 and the other 36 pictures in Trial Blocks 4-6 (the second half-experiment). In each trial block, familiarization included 18 singletons and 6 triplets. This was followed by 4 list recall trials in random order: two lists of nine singletons, and two lists of three triplets. The singletons and triplets were the same as in the familiarization session, but the orders of singletons and triplets within each list varied randomly. These orders changed from one block to the next within the half-experiment, while the particular triplets remained the same.

In the rehearsal suppression group, adult participants were trained to say the word *the* twice per second during the picture presentation sequence in list recall. On each trial, when they said it the second time, the experimenter initiated the picture sequence with a key press and this rehearsal suppression was to continue until the picture sequence ended. There was a practice sequence using pictures that were not used in the experiment proper and practice was repeated until the participant reliably said *the* at the correct rate; each trial was monitored by the experimenter, and reminders and further training were given between trials on the rare occasions when they were necessary.

One might wonder whether a more complex suppression task could be more effective. There is a somewhat larger effect of suppression tasks when the item to be articulated keeps changing throughout the procedure (Cowan, Cartwright, Winterowd, & Sherk, 1987). However, the problem with such a task is that the additional effect could come from item interference (Oberauer & Lewandowsky, 2008) or distraction (Gavens & Barrouillet, 2004) rather than better prevention of rehearsal. With repetition of the word *the* as the suppression task, a previous study showed that adults' performance was reduced to a level and form comparable to children in the early elementary years (Cowan, Saults et al., 2006). Similar results were obtained with repetition of a single letter (Cowan et al., 1987). Repetition of *the* was also sufficient to remove the phonological length contribution to immediate recall, leaving behind a fixed capacity of about 3 chunks regardless of chunk length or list length (Chen & Cowan, 2009). It is therefore well-suited to examine performance in adults without a rehearsal advantage over young children.

Results

Benefits of set familiarization. For the purpose of the analyses, because familiarization involved triplets, Set 1 was defined as Items 1-3 in the list to be recalled, Set 2 was defined as Items 4-6, and Set 3 was defined as Items 7-9. An ANOVA on the mean access across the sets was conducted with three groups (children, adults carrying out rehearsal suppression, and adults without suppression) and with the familiarization condition (distinguishing lists of nine singletons from lists of three triplets) as a within-participant factor. As shown in Figure 5, there was an overall group effect $F(2,69)=52.34$, $MSE=0.01$, $\eta_p^2=.60$, with access in the children

(M=.70, SEM=0.01) below both of the two adult groups, which did not differ (suppression, M=.88, SEM=0.01; no suppression, M=.90, SEM=0.01), according to Newman-Keuls pairwise tests. There was also an overall effect of condition, F(1,69)=5.88, MSE=0.01, η_p^2=.08, demonstrating a small benefit of set familiarization, with a higher proportion of sets accessed for triplets than for singletons (see Figure 5). The interaction of these factors did not approach significance.

- - - - Figure 5 about here - - - -

The same effects were significant for the completion of accessed sets. As shown in Figure 6, there was again a large effect of group, F(2,69)=37.39, MSE=0.01, η_p^2=.52, with Newman-Keuls tests again showing that the children had much lower completion rates than either group of adults, which did not differ. One difference is that, whereas the effect of set familiarization on access was small, the effect of set familiarization on completion was huge, F(1,69)=150.97, MSE=0.01, η_p^2=.67, with a higher proportion of completion of sets for triplets than for singletons. This is as it should be, inasmuch as the set familiarization primarily allows better completion by strengthening inter-item associations in the set. Finally, the interaction of group and familiarization type did not approach significance (F<1), indicating as in Experiments 1 and 2 that children benefited from familiarization about as much as adults did.

Although rehearsal suppression did not have much of an effect on the overall level of performance in adults, it did alter the serial position functions for set completion. For that measure, the three-way interaction of rehearsal group, familiarization condition, and serial position was significant, F(2,92)=4.69, MSE=0.01, η_p^2=.09. In the group allowed to rehearse, completion scores across the three item-pair positions fell off much more slowly for triples (M=.95, .86, & .73) than for singletons (M=.88, .67, & .44). In contrast, in the group carrying out rehearsal suppression, the fall-off was similar for triples (M=.93, .86, & .68) and singletons (M=.82, .59, & .46). There was no comparable interaction effect for pair access. This pattern of results suggests that rehearsal might have been used in one adult group to maintain the associative information between items within a triplet as recall progressed.

- - - - Figure 6 about here - - - -

Chunk capacity. Once more, chunk capacity was examined by an analysis of access in triplets with completion as a covariate. This ANCOVA analysis produced not only an effect of the covariate, $F(1,68)=15.14$, $MSE=0.01$, $\eta_p^2=.18$, but also an effect of group, $F(2,68)=22.75$, $MSE=0.01$, $\eta_p^2=.40$. Thus, the covariate cannot account for the simple effect of group on access scores. Newman-Keuls tests showed all three groups to differ from one another in the proportion of sets accessed: adults with no suppression (M=.93), adults under rehearsal suppression (M=.89), and children (M=.70).

Discussion

This third experiment overcame the ceiling effect in set access observed in Experiment 2 through the presentation of 9-item lists, and showed that there was an age difference in set access, even with set completion controlled.

Another interesting aspect of this experiment is that there was only a small effect of rehearsal suppression in adults on access or completion of sets. This finding with pictures to be recalled is in stark contrast to the usual finding of a large effect of rehearsal suppression on recall of verbal materials, making adults under suppression comparable to elementary school children without it (e.g., Cowan et al., 2006). The small overall size of the effect of suppression may occur because of the long list length and/or the use of pictorial materials with inter-item associations. With list lengths longer than the capability of the phonological loop, individuals tend to switch to a mode of maintenance that does not depend on phonological rehearsal (Baddeley & Larsen, 2007). Maintenance of non-verbal materials often is found to be impervious to articulatory suppression. For example, Morey and Cowan (2004) found that memory for a visual array was not impeded by repetition of a known 7-digit telephone number, though it was impeded by repetition of a random 7-digit memory load.

Across-Experiment Analyses

Set Access After Matching on Set Completion

By combining results across experiments it was possible to examine, in another way, set access in list recall independent of set completion. We did this by matching across age groups on set completion. The results for the learned-association (familiarized pair and triplet) conditions were averaged for each participant. Table 3 shows mean access rates in list recall, separated according to the mean set completion rounded off to the nearest 0.1. The table shows that there were enough participants to compare adults with no rehearsal suppression to children, using only participants with rounded completion scores of 1.0, 0.9, and 0.8. An ANOVA of set access included age group and the rounded completion score as between-participant variables and produced not only an effect of the rounded completion score, $F(2,113)=6.33$, $MSE=0.01$, $\eta_p^2=.10$, but also an effect of age group, $F(1,113)=20.35$, $MSE=0.01$, $\eta_p^2=.15$ (see Table 3). Further validating this analysis, a fine-grained examination of the precise completion rates for the participants included in the analysis showed no effect of age on actual completion rates within each rounded-off category. (The means were 1.0, .89, and .81 for children and .99, .90, and .80 for adults for the three rounded-off categories, respectively.) The analyses are consistent with the conclusion from all three experiments that chunk capacity increased across age groups.

- - - - Table 3 about here - - - -

Fair Estimates of Capacity

The matching procedure also can be used to compare capacity in children and adults. If one assumes that, in the familiarized pair and triplet conditions, each presented set is a single chunk, then the capacity estimate is the proportion of sets accessed, multiplied by the number of sets presented. That assumption, however, is shaky when the completion rate is low because, for that situation, chunks presumably did not always form and therefore each presented set may require, on average, more than one slot in working memory (Chen & Cowan, 2005, 2009). We therefore restricted the analysis of capacity to the participants used in the matching analysis. In calculating a single capacity estimate for each participant (in the manner just described), we restricted the data to the familiarized pair conditions in Experiments 1 and 2, and the familiarized triplet condition in Experiment 3; these are the conditions in which at least 3 familiarized sets were presented in the list. There was an effect of age group, $F(1,113)=14.33$, $MSE=0.12$, $\eta_p^2=.11$. The capacity for 69 adults was 2.93 (SEM=0.05), and the capacity for 50 children was lower at 2.66 (SEM=0.05). There was no effect of completion within the high range used (0.8-1.0), $F(2,113)=1.17$, $p>.1$, and the age by completion interaction was not significant, $F(2,113)=2.17$, $p>.1$. These are therefore fairly stable estimates, though they are available for high-performing participants only. From Table 2, one can estimate the capacity for all participants if one is willing to disregard the possible effect of completion rates, by multiplying the set-access rate by the number of sets in the list for familiarized-pair and familiarized-triplet conditions; the estimates are a bit lower and exaggerate the age effect further.

General Discussion

An important but difficult question in child development is whether improvements in list recall result from more chunks recalled, larger chunks recalled, or both. The current studies contribute to the resolution of this issue by coding recall in terms of its sub-structure and not just overall accuracy. In all three experiments, we have provided a confirmation of a bold conclusion that has been drawn before only rarely: that the developmental increase in working memory span occurs at least partly because of a growth in the number of chunks that can be retained concurrently. Specifically, in each experiment, associations between sets of words were taught and lists were composed of sets that did or did not include these associations. It was found that adults accessed more sets within a list than did children in the early elementary school years, where access was defined as recall of at least one word from the set. Further, it was found that this difference in access to word sets persisted even when the proportion of items recalled from each accessed set, i.e. set completion, was controlled across age groups. We found this for lists of spoken words in Experiment 1 and for lists of pictures of common nouns in Experiments 2 and 3. Using completion as a proxy for the level of association between words used in recall, and controlling for completion across age groups, access provides a good estimate of how many independent, unassociated elements were recalled. It suggests that capacity increases with age.

One way in which this conclusion could be mistaken is if adults were better than children at combining the chunks that they studied to form super-chunks (cf. Ericsson, Chase, & Faloon, 1980). For example, suppose that a participant learned the associations "brick-fish" and "dog-ball." If these pairs occurred one after the other in a list, the participant might be able to associate them rapidly during the list presentation, to form a single, longer super-chunk, "brick-fish/dog-ball." If this new conglomerate occupied only a single slot in working memory, it would free up space for more additional items. Given that our measure of capacity does not take this into account, the result would be an increase in the observed set access in the studied-association conditions compared to the other conditions. In contrast to that kind of prediction, though, adults showed no difference between familiarization conditions in set access in any of the experiments (Figures 1, 3, & 5). Instead, familiarization conditions influenced only set completion in adults (Figures 2, 4, & 6). For this reason, the better access of adults compared to children cannot be explained in terms of the adults forming super-chunks. A developmental increase in chunk capacity is the only way we have found to explain the results.

The finding that children did have higher access scores for pairs and triplets than for singletons suggests that they were unable to form on-line chunks as well as adults. That is also evident in the age group differences in completion scores and in the interaction of familiarization conditions with age group in Experiment 1, in which only the children accessed more sets with pair training than without it. Given the age differences in set access with set completion as a covariate or matched across groups, however, it seems unlikely that this ability can account for age differences in set access. It is even clear from the graphs that set access for children in the pair or triplet conditions is still not as high as it is for adults in the singleton conditions.

The conclusion that capacity increases with development in childhood has been drawn by some previous investigators (e.g., Burtis, 1982; Case, 1972, 1995; Cowan et al., 1999; Pascual-Leone, 1970, 2005) but the context of the conclusion was limited. Most of the previous work did not include a measure of chunk integrity. Most of the studies assumed that each item was a separate chunk, and therefore there was no control for the possibility that more mature individuals might be more likely to form associations on line that would reduce the number of items that had to be accessed independently to be recalled. Burtis (1982) produced stimuli that were highly likely to be combined into a single chunk but that is not quite the same as measuring an aspect of chunking, as the completion measure does. The difficulty in drawing such a conclusion is that it is difficult to rule out the alternative possibility of more efficient use of a constant number of slots in working memory (e.g., Case et al., 1982; Guttentag, 1984; Kail, 1997; but see Towse & Hitch, 1995).

Gilchrist et al. (2009) did provide a measure of set completion, in the form of the completion of simple English sentences that were presented as list items. That study showed that adults accessed more sentences than children, even though both groups completed about 80% of the accessed sentences on average. The present study extends this logic, and the conclusion of an age increase in the number of chunks accessed in working memory, to a situation in which the list items were taught during the experiment rather than pre-learned, and with lists closer to those used in a vast number of previous studies of recall.

Capacity measures in the present study can be compared to the benchmark results of Chen and Cowan (2009). They taught adults pairs of words to a 100% correct cued recall criterion in order to be able to say that each pair was a strong chunk in working memory. They also included rehearsal suppression, and scored the number of items correct without regard to item order as in the present study. There were lists of 4, 6, 8, and 12 singletons and lists of 4 or 6 learned pairs. The finding was that individuals accessed about 3 singletons in the singleton conditions and about 3 learned pairs in the paired conditions. This estimate of 3 chunks is not far from the adult estimate of 2.93 offered above, which is higher than the estimate of 2.66 for children with comparably high set-completion rates. These are slight underestimates inasmuch as the completion rate was still mostly below 1.0, but they otherwise are in reasonably good accord with previous studies using different techniques (e.g., Cowan et al., 1999, 2005; Gilchrist et al., 2009).

A striking aspect of the present Experiment 3 was how little difference rehearsal suppression made in the recall of picture sets (e.g., see Figures 5 & 6). Previous studies suggest that, in contrast, rehearsal suppression during memory for verbal-spatial associations has a profound effect on adults' recall, making their performance rather similar to elementary school children (e.g., Cowan et al., 2006). The general non-significance of the rehearsal suppression factor in our Experiment 3 suggests that participants did not rely much on verbal labeling in their recall responses. It may be that the associative structure drove them to think more in visual and semantic terms in order to capitalize upon these associations, whereas rehearsal is a more important factor when the strategy is more heavily reliant on phonological information (cf. Baddeley, 1986).

The previous literature (e.g., Cole, Frankel, & Sharp, 1971; Moely, Olson, Halwes, & Flavell, 1969) suggests that children can use grouping information adequately in free recall provided that there is adequate support for this grouping. It has been harder to provide evidence of grouping in children in serial recall (Cowan, Elliott, et al., 2006; Dempster, 1978; Towse et al., 1999; though see Harris & Burke, 1972). The present work does suggest a developmental increase in the number of items that can be bound together to form a chunk. In Experiment 1, this emerged as a developmental increase in the likelihood that a learned word pair that was accessed (i.e., from which at least one word was recalled) would be completed (i.e., from which both words were recalled). In Experiments 2 and 3, which used a picture-presentation procedure and learned triplets as well as pairs, the developmental increase in the proportion completion of a chunk was replicated.

An important additional question is whether the age difference in chunk completion occurred because the long-term memory familiarization was less successful in children. That does not appear to be the case, given that the age difference in pair or triplet completion did not depend on the familiarization condition; in all three experiments, both age groups benefited from familiarization to a similar extent, and the potent effect of familiarization on set completion did not interact with age at all. Thus, the age differences we have observed appear to come largely or completely from working memory differences, and not from differences in long-term learning.

In sum, we have demonstrated a new method to distinguish between associative knowledge and item capacity in short-term recall, and have shown that this method supports the hypothesis that there is an increase in capacity with development in childhood. Recent research also has examined behavioral consequences of capacity limits in a nonverbal domain (e.g., Gobet & Clarkson, 2004), brain signatures of verbal short-term memory limits (e.g., Talmi, Grady, Goshen-Gottstein, & Moscovitch, 2005), brain signatures of capacity limits for objects and complexity (e.g., Xu & Chun, 2006), and decreases in chunk capacity with human aging (Naveh-Benjamin et al., 2007). These provide interesting grounds for interdisciplinary work in the near future, the better to understand working memory capacity, a basic human cognitive resource that has intrigued researchers for as long as the field of human experimental psychology has existed.

References

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

Baddeley, A.D., & Larsen, J.D. (2007). The phonological loop: Some answers and some questions. *The Quarterly Journal of Experimental Psychology. 60*, 512-518.

Binet, A., & Simon, T. (1916/1980). *The Development of Intelligence in Children,* Williams Printing Company, Nashville, TN.

Bolton, T.L. (1892). The growth of memory in school children. *American Journal of Psychology*, *4*, 362-380.

Burtis, P.J. (1982). Capacity increase and chunking in the development of short-term memory. *Journal of Experimental Child Psychology, 34*, 387-413.

Case, R. (1995). Capacity-based explanations of working memory growth: A brief history and reevaluation. In F.E. Weinert & W. Schneider (eds.), *Memory performance and competencies: Issues in growth and developmen*t. Mahwah, NJ: Erlbaum. (pp. 23-44)

Chen, Z., & Cowan, N. (2005). Chunk limits and length limits in immediate recall: A reconciliation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1235-1249.

Chen, Z., & Cowan, N. (2009). Core verbal working-memory capacity: The limit in words retained without covert articulation. *The Quarterly Journal of Experimental Psychology,* 62, 1420–1429.

Chi, M.T. (1976). Short-term memory limitations in children: Capacity or processing deficits? *Memory & Cognition, 4*, 559-572.

Chi, M.T.H. (1978). Knowledge structures and memory development. In R. Siegler (Ed.), *Children's thinking: What develops*? Hillsdale, NJ: Erlbaum.

Cole, M., Frankel, F., & Sharp, D. (1971). Development of free recall learning in children. *Developmental Psychology, 4*, 109-123.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87-185.

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

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

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

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

Cowan, N., Nugent, L.D., Elliott, E.M., Ponomarev, I., & Saults, J.S. (1999). The role of attention in the development of short-term memory: Age differences in the verbal span of apprehension. *Child Development, 70*, 1082-1097.

Cowan, N., Saults, J.S., Elliott, E.M., & Moreno, M. (2002). Deconfounding serial recall. *Journal of Memory and Language*, *46*, 153-177.

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

Dempster, F.N. (1978). Memory span and short-term memory capacity: A developmental study. *Journal of Experimental Child Psychology, 26*, 419-431.

Ericsson, K.A., Chase, W.G., & Faloon, S. (1980). Acquisition of a memory skill. *Science, 208*, 1181-1182.

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

Fearnley, S. (1997). MRC Psycholinguistic Database search program. *Behavior Research Methods, Instruments, & Computers, 29*, 291-295.

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.

Frankish, C. (1985). Modality-specific grouping effects in short-term memory. *Journal of Memory and Language, 24*, 200-209.

Gavens, N. and P. Barrouillet (2004). Delays of retention, processing efficiency, and attentional resources in working memory span development. *Journal of Memory & Language, 51*, 644-657.

Gilchrist, A.L., Cowan, N., & Naveh-Benjamin, M. (2009). Investigating the childhood development of working memory using sentences: New evidence for the growth of chunk capacity. *Journal of Experimental Child Psychology, 104*, 252-265.

Gobet, F., & Clarkson, G. (2004). Chunks in expert memory: Evidence for the magical number four… or is it two? *Memory, 12*, 732-747.

Guttentag, R. E. (1984). The mental effort requirement of cumulative rehearsal: A developmental study. *Journal of Experimental Child Psychology, 37*, 92-106.

Harris, G.J. & Burke, D. (1972). The effects of grouping on short-term serial recall of digits by children: Developmental trends. *Child Development, 43*, 710-716.

Hitch, G. J., Halliday, M. S., & Littler, J. E. (1989). Item identification time and rehearsal rate as predictors of memory span in children. *The Quarterly Journal of Experimental Psychology, 41A*, 321-337.

Hulme, C., & Tordoff, V. (1989). Working memory development: The effects of speech rate, word length, and acoustic similarity on serial recall. *Journal of Experimental Child Psychology, 47,* 72-87.

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

Jacobs, J. (1887). Experiments on "prehension." *Mind, 12*, 75-79.

Kail, R. (1997). Phonological skill and articulation time independently contribute to the development of memory span. *Journal of Experimental Child Psychology, 67*, 57-68.

Marmurek, H.H., & Johnson, N.F. (1978). Hierarchical organization as a determinant of sequential learning. *Memory & Cognition, 6*, 240-245.

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

Miller, Gregory A., & Chapman, J.P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology, 110*, 40-48.

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

Moely, B.E., Olson, F.A., Halwes, T.G., & Flavell, J.H. (1969). Production deficiency in young children's clustered recall. *Developmental Psychology, 1*, 26-34.

Naveh-Benjamin, M., Cowan, N., Kilb, A., & Chen, Z. (2007). Age-related differences in immediate serial recall: Dissociating chunk formation and capacity. *Memory & Cognition, 35*, 724-737.

Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: decay, temporal distinctiveness, or interference? *Psychological Review, 115*, 544-576.

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

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

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

Ramsay, M.C., & Reynolds, C.R. (1995). Separate digit tests: A brief history, a literature review, and a reexamination of the factor structure of the test of memory and learning (TOMAL). *Neuropsychology Review, 5*, 151-171.

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object set: The role of surface detail in basic-level object recognition. *Perception, 33*, 217-236.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools Inc.

Slak, S. (1970). Phonemic recoding of digital information. *Journal of Experimental Psychology, 86*, 398-406.

Talmi, D., Grady, C.L., Goshen-Gottstein, Y., & Moscovitch, M. (2005). Neuroimaging the serial position curve: A test of single-store versus dual-store models. *Psychological Science, 16*, 716-723.

Thorn, A.S.C., Gathercole, S.E., & Frankish, C.R. (2005). Redintegration and the benefits of long-term knowledge in verbal short-term memory: An evaluation of Schweickert’s (1993) multinomial processing tree model. *Cognitive Psychology, 50*, 133–158.

Towse, J. N., & Hitch, G. J. (1995). Is there a relationship between task demand and storage space in tests of working memory capacity? *Quarterly Journal of Experimental Psychology, 48A*(1), 108-124.

Towse, J.N., Hitch, G.J., & Skeates, S. (1999). Developmental sensitivity to temporal grouping effects in short-term memory. *International Journal of Behavioral Development, 23*, 391-411.

Tulving, E., & Patkau, J.E. (1962). Concurrent effects of contextual constraint and word frequency on immediate recall and learning of verbal material. *Canadian Journal of Psychology, 16*, 83-95.

Wechsler, D. (1991). *Manual for the Wechsler Intelligence Scale for Children--Third Edition* (WISC-III). The Psychological Corporation, San Antonio, TX.

Wundt, W. M. (1906). *Einfiihrung in die psychologie.* R. Voigtländer, Leipzig.

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

Author Note

This work was funded by NIH grant R01 HD-21338. We thank Vanessa Evoh, Ooshene Fox, Sarah Moore, Chris Blume, Georgette Richardson, and Sushila Pepe for assistance.

Address correspondence to Nelson Cowan, Department of Psychological Sciences, University of Missouri, 18 McAlester Hall, Columbia, MO 65211. E-mail CowanN@missouri.edu.

Footnote

[1] The proportions correct by serial position according to either strict serial position scoring or free scoring were unsurprising; we consider them tangential to the present aim inasmuch as they can illuminate only the selection of items for recall, and not the capacity per se.

Table 1

Procedural Details of Each Experiment

Detail	Experiment 1	Experiments 2 and 3
Stimuli	Spoken Words	Pictures
Familiarization Conditions	Singletons, Pairs (1 or 2 words in a row)	Singletons, Pairs[a], Triplets (1, 2, or 3 pictures / exposure)
Familiarization Task	Repeat each word or pair	Name each picture
Cued Recall Test?	Yes (and identify singletons as such)	No
Serial Recall Task	Listen to list, then repeat it	Watch list of pictures, then repeat their names
Serial Recall Conditions	Novel Words, Famil. Singletons, Famil. Pairs (1 type per list)	Famil. Singletons, Pairs[a], Triplets (1 type per list)
Serial Recall Presentation	Spoken words, one at a time Always grouped into pairs	Pictures, one at a time Even pacing
Serial Recall List Length	6 words or 8 words	Experiment 2: 6 pictures Experiment 3: 9 pictures
Serial Recall Response Mode	Oral	Oral

Note. All experiments included children and adults but, in Experiment 3 only, a second group of adults carried out articulatory suppression during the serial recall procedure.

[a] Pairs of pictures were included in Experiment 2 but not in Experiment 3.

Table 2

Access and completion rates for each age group in familiarized pair and triplet conditions (with SEM)

		Pairs		Triplets	
Experiment/Group/List Length	N	Access	Completion	Access	Completion
1/Children/6-item	30	.89(.03)	.89(.02)	---	---
1/Adults/6-item	30	.94(.03)	.95(.02)	---	---
1/Children/8-item	30	.68(.04)	.94(.02)	---	---
1/Adults/8-item	30	.78(.04)	.93(.02)	---	---
2/Children/6-item	21	.87(.02)	.85(.02)	.95(.02)	.79(.02)
2/Adults/6-item	21	.98(.02)	.94(.02)	.99(.02)	.91(.02)
3/Children/9-item	24	---	---	.70(.01)	.64(.02)
3/Adults/9-item	24	---	---	.93(.01)	.85(.02)
3/Adults-suppression/9-item	24	---	---	.89(.01)	.82(.02)

Note. Access refers to the recall of at least one word from a set and completion refers to the proportion of words recalled from accessed sets. Here the sets were pairs or triplets presented to participants in a familiarization phase of each experiment, before they were included in lists to be recalled. Suppression refers to rehearsal suppression (repetition of the word *the*) during the recall task.

Table 3

Proportion of Sets Accessed as a Function of Set Completion in List Recall, All Experiments

	Proportion Completion of Accessed Sets, Rounded to Nearest 0.1						
Statistic	1.0	0.9	0.8	0.7	0.6	0.5	0.4
Adults, No Suppression (all experiments)							
Mean	0.90	0.94	0.92	0.86	0.71	--	--
SEM	0.02	0.01	0.03	0.02	--	--	--
N	28	32	9	5	1	--	--
Adults, Suppression (Experiment 3)							
Mean	--	0.92	0.89	0.89	0.77	--	--
SEM	--	0.01	0.03	0.03	0.03	--	--
N	--	11	8	2	3	--	--
Children (all experiments)							
Mean	0.76	0.89	0.82	0.78	0.73	0.67	0.58
SEM	0.04	0.03	0.03	0.03	0.03	0.01	0.00
N	14	17	19	9	10	4	2

Note. Access refers to the recall of at least one word from a set and completion refers to the proportion of words recalled from accessed sets. Table includes only the conditions in which participants studied associations between items that were then included in the list to be recalled as learned pairs or triplets. Suppression refers to rehearsal suppression during the recall task.

Figure Captions

Figure 1. Access to sets (defined as the proportion of two-item sets from which at least one word was recalled) for each age group and condition in Experiment 1. Famil. = familiarized before the words were used in recall, as singletons or associated pairs. Novel = not familiarized. Error bars are standard errors.

Figure 2. Rate of completion of sets (defined as the proportion of items recalled from accessed two-item sets) for each age group and condition in Experiment 1. Famil. = familiarized before the words were used in recall, as singletons or associated pairs. Novel = not familiarized. Error bars are standard errors.

Figure 3. Access to sets (defined as the proportion of two- or three-item sets from which at least one word was recalled) for each age group and condition in Experiment 2. Sngltn = singleton during familiarization, cntl = control. The familiarized singletons were scored as two- or three-item sets as controls for comparison with familiarized pairs or triplets. Error bars are standard errors.

Figure 4. Rate of completion of sets (defined as the proportion of items recalled from accessed sets) for each age group and condition in Experiment 2. Sngltn = singleton during familiarization, cntl = control. The familiarized singletons were scored as two- or three-item sets as controls for comparison with familiarized pairs or triplets. Error bars are standard errors.

Figure 5. Access to sets (defined as the proportion of three-item sets from which at least one word was recalled) for each group and condition in Experiment 3. Singletons were examined in three-item sets as a control for triplets. Error bars are standard errors.

Figure 6. Rate of completion of sets (defined as the proportion of items recalled from

accessed three-item sets) for group and condition in Experiment 3. Singletons were examined in three-item sets as a control for triplets. Error bars are standard errors.

Figure 1

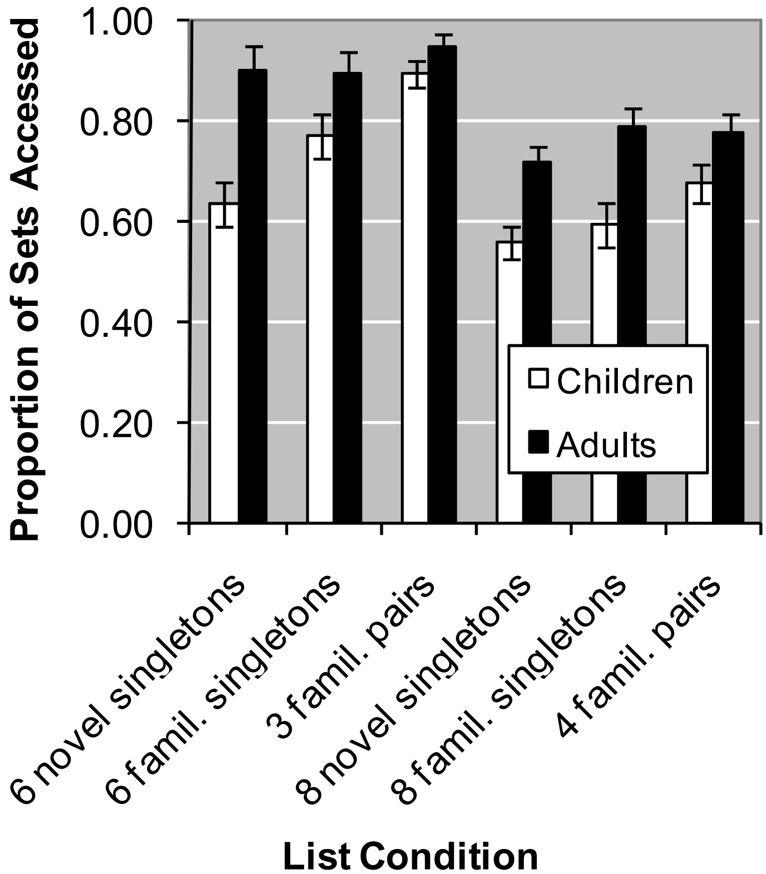

1.00
0.80
0.60
0.40
0.20
0.00
Proportion of Sets Accessed
Children
Adults
6 novel singletons
6 famil. singletons
3 famil. pairs
8 novel singletons
8 famil. singletons
4 famil. pairs
List Condition

Figure 2

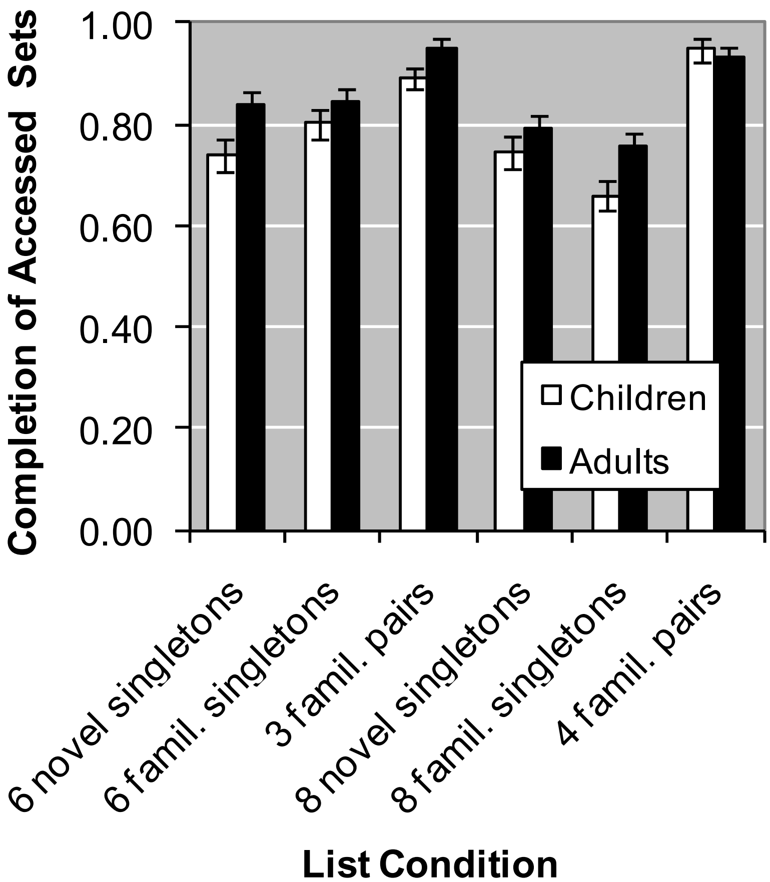

Completion of Accessed Sets
1.00
0.80
0.60
0.40
0.20
0.00
Children
Adults
6 novel singletons
6 famil. singletons
3 famil. pairs
8 novel singletons
8 famil. singletons
4 famil. pairs
List Condition

Figure 3

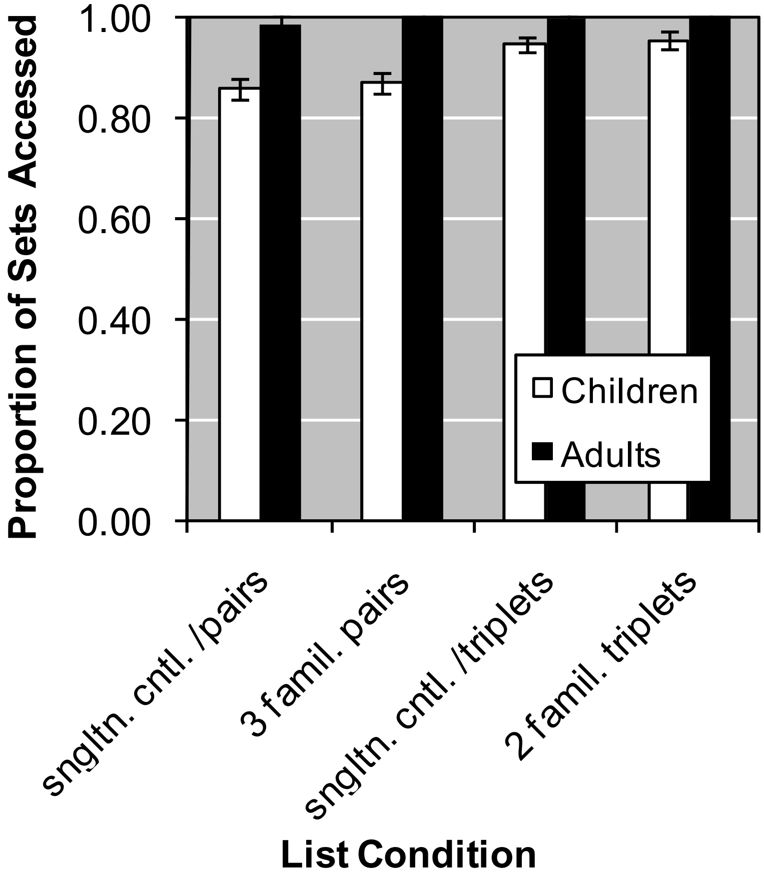

1.00
0.80
0.60
0.40
0.20
0.00
Proportion of Sets Accessed
Children
Adults
sngltn. cntl. /pairs
3 famil. pairs
sngltn. cntl. /triplets
2 famil. triplets
List Condition

Figure 4

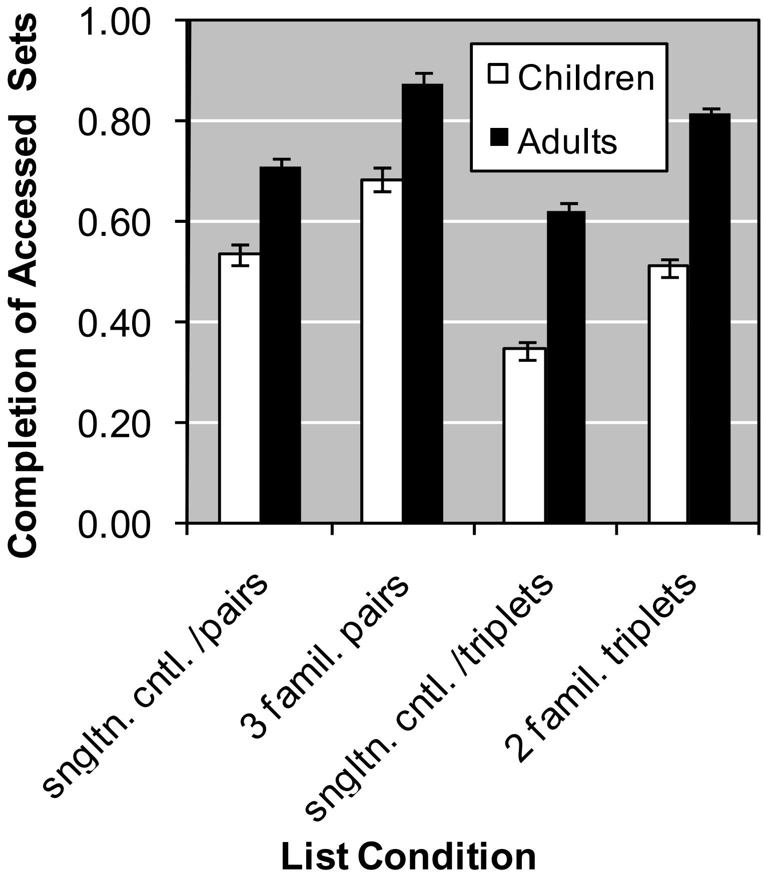

Completion of Accessed Sets
1.00
0.80
0.60
0.40
0.20
0.00
Children
Adults
sngltn. cntl. /pairs
3 famil. pairs
sngltn. cntl. /triplets
2 famil. triplets
List Condition

Figure 5

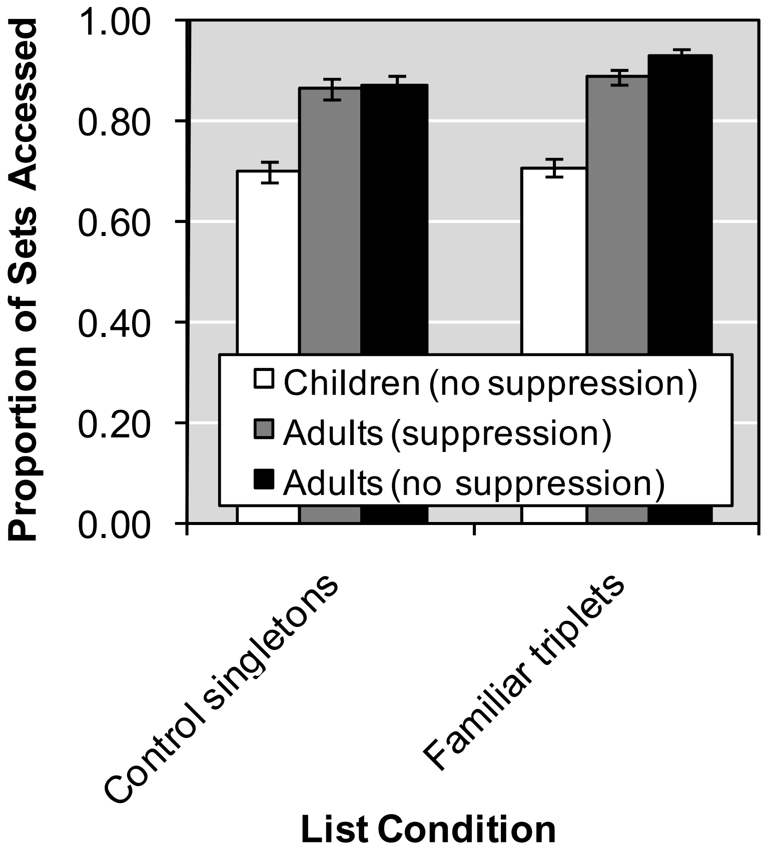
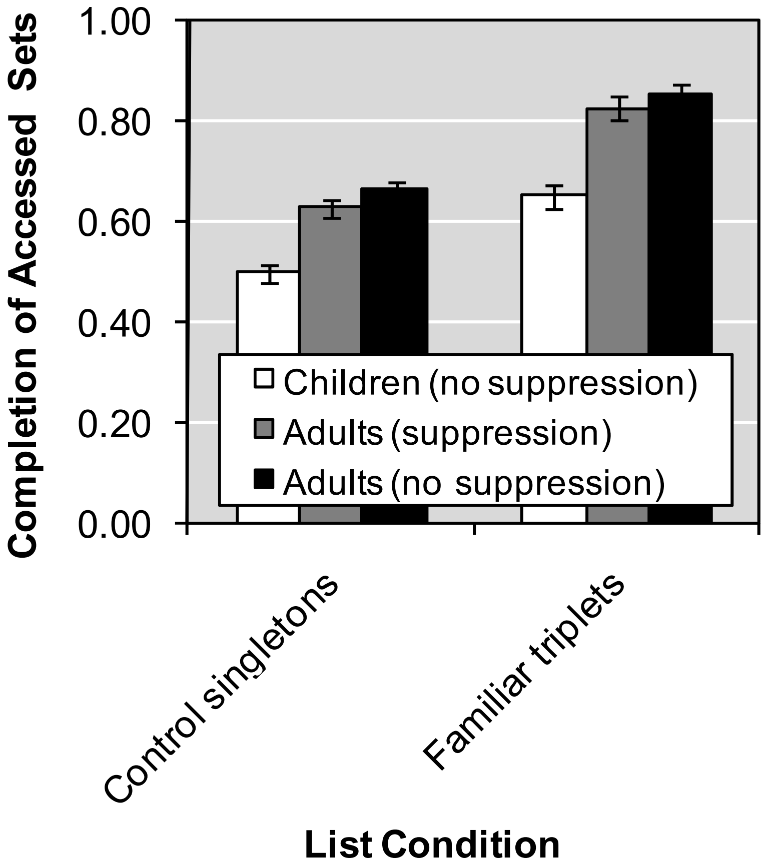

Proportion of Sets Accessed
1.00
0.80
0.60
0.40
0.20
0.00
Children (no suppression)
Adults (suppression)
Adults (no suppression)
Control singletons
Familiar triplets
List Condition

Figure 6


Completion of Accessed Sets
1.00
0.80
0.60
0.40
0.20
0.00
Children (no suppression)
Adults (suppression)
Adults (no suppression)
Control singletons
Familiar triplets
List Condition