Running head: SUPPLEMENT FOR "CORE ASSUMPTIONS OF AGE-LINKED RESOURCE CAPACITY THEORIES"

Supplemental Information for "The Core Assumptions of Age-linked Resource-capacity Theories":

Statistics, Participant Information, Tables of Results and Stimulus Appendix.

Christopher B. Hadley

&

Donald G. MacKay

Psychology Department

University of California, Los Angeles

Los Angeles, CA 90024-1563

Phone: 310-825-8465

Corresponding author: cbhadley@gmail.com

Supplemental Information for "The Core Assumptions of Age-linked Resource-capacity Theories":

Statistics, Participant Information, Tables of Results and Stimulus Appendix.

This paper provides supplemental information for two experiments that test the predictive adequacy of SCART, a simulation described in MacKay, Hadley and Abrams (submitted) of the core assumptions of resource-capacity theories of immediate memory and cognitive aging. For introductory information, the preconditions for testing SCART, and the mathematical derivation of SCART predictions, together with theoretical implications of the results, see MacKay et al. The present text closely matches MacKay et al. except for insertion of detailed participant information, statistics, stimuli, and tables of results. We also reproduce renumbered versions of Figures 2-12 from MacKay et al. for ease of following the text.

Experiment 1- Rapid Presentation Rates

For Experiment 1, the phrase and no-phrase lists were presented at 170/200 ms/word for young adults and 200/270 ms/word for older adults. We adopted these relatively rapid rates to rule out rehearsal strategies and we adopted one shared and one different presentation rate for young and older participants for four reasons: to avoid floor effects for older adults at the fastest rate (170ms/word), to avoid ceiling effects for young adults at the slowest rate (270ms/word), to enable comparison of the identical shared rate for young versus older adults (200ms/word), and to ensure that both age groups could readily perceive the words (see also MacKay & Miller, 1994). SCART predicted exact levels of recall for groups with mean capacities of 3, 4 and 5.5 chunks (see the Appendix). SCART also predicted a main effect of age, with better recall for young than older adults; better recall of extraneous words in phrase than no-phrase lists, especially for older adults; better recall of extraneous words in longer than shorter lists, especially for young adults; and no difference in mean number of chunks recalled for young or older adults as a function of list-length and list-type (phrase versus no-phrase).

Method

Participants

Participants were 30 UCLA undergraduates who received course credit for participation (mean age = 19.4 years, SD= 2.32 years) and 31 healthy, community-dwelling older participants (mean age = 70.5 years, SD = 4.46 years) from the UCLA Cognition and Aging Laboratory participant pool who received $10.00/hour.

Participants first completed the Nelson-Denny vocabulary test and the forward and backward digit span tests, and answered a questionnaire concerning their health, education and other background characteristics, with results shown in Table 1. All participants were fluent speakers of English who reported normal or corrected-to-normal vision.

Materials

Materials were 32 lists containing 6 or 7 high frequency words (mean frequency = 445 in Kucera & Francis, 1968). Each list came in a phrase and and no-phrase version. No-phrase versions consisted exclusively of phonological dissimilar words that were unrelated in semantics and syntax and were not part of recognizable phrases. Phrase lists were identical to no-phrase lists, except for the substitution of two words immediately before or after the two critical words in no-phrase lists. These word substitutions formed two highly familiar phrases with the critical words in phrase lists, and the critical words enabled comparisons of recall for identical words in identical positions in chunks in phrase lists versus not in chunks in no-phrase lists (see Appendix). List-position of the critical words varied systematically across lists, and phrase and no-phrase lists were presented in 8 different randomized sequences with 16 phrase and 16 no-phrase versions in each sequence. Each participant saw all 32 lists at rates divided equally between 170 and 200 ms/word for young participants or 200 and 270 ms/word for older participants.

Procedure

Instructions indicated that participants would see word lists presented briefly one word at a time centered on the computer screen, and that they were to recall aloud as many words as possible in order if possible as soon as the list ended. If they were uncertain of the original list-order of a word, participants were to recall it in any order. The instructions also indicated that the lists were mostly unrelated words, but that adjacent words would sometimes form phrases that may help recall. A series of question marks (“?????????”) centered on the screen indicated list end and called for immediate recall

A Macintosh computer running Psyscope (Cohen, MacWhinney, Flatt & Provost, 1993) presented the words via RSVP at the specified rate per word with no visual cues as to which words were phrase. Before the experiment proper, each participant saw four representative practice lists, two at each rate. Trials were self-paced; the participant initiated each trial by pressing the space bar, after which a 2 s warning (“Get ready for the next list”) and 1.2 s blank screen preceded the next list. The experimenter scored the responses and order of recall on pre-printed answer sheets, and wrote any intrusions (words recalled that were not part of the presented list) in their relative position in the list. In addition, each session was tape recorded to allow scoring cross-checks.

Results and Discussion

Results for Experiment 1 strongly supported all four preconditions for valid tests of SCART (see MacKay, Hadley, & Abrams, *submitted*, for details). We therefore present our main results before our precondition results to facilitate exposition. All results reported as significant reached $p < .05$. Unless otherwise noted, all analyses ignored the serial order of recall.

Main Results

Table 2 shows the mean number of chunks correctly recalled by age group and list type for all lists in Experiment 1. Significantly fewer chunks were recalled in phrase than no-phrase lists for both young adults, $t = (29) = -18.456$, $p < .001$, and older adults, $t = (30) = -11.778$, $p < .001$, and the overall mean chunks recalled in no-phrase lists was about 3 chunks for older adults, and about 4 chunks for young adults, approximating the 4-chunk limit that Cowan (2001) postulated for immediate memory tasks. We therefore initially compared the data for older adults with capacity-3 predictions of SCART and with capacity-4 predictions for young adults. Because SCART predicted measurably larger interactions in our data for capacity-5 or greater (see Figures 2-4), we initially planned to analyze separately the data for young participants with mean recall of 5 chunks or greater to determine whether these interactions emerged with larger capacities. However, because no young participants had mean recall of 5 chunks or greater, we performed a median split of the data for young participants based on mean capacity per participant (determined by performance for the no-phrase lists) and conducted additional analyses with capacity as a between-subjects factor.

Freed-capacity results. Figure 1 shows the mean recall of extraneous words by age and list-type (right panel), together with predicted recall in SCART (the freed-capacity predictions; left panel) for capacity-3, -4, and 5.5 (-5 and -6 combined): better recall of extraneous words in phrase than no-phrase lists, albeit less so for young (capacity-4 and above) adults; and a main effect of age, with better recall for young than older adults. A 2 (age: young vs. older) x 2 (list type: phrase vs. No-phrase) mixed ANOVA on the recall data indicated no main effect of age, $F(1, 59) = 1.316$, $MSE = .029$, $p = .256$, and no age x list type interaction, $F(1, 59) = 2.415$, $MSE = .011$, $p = .125$, and no main effect of list type, $F(1, 59) = .002$, $MSE = .011$, $p = .966$. These null effects of list-type were not due to insufficient power to detect a difference. Power analysis indicated a power of .873 to detect the large effect predicted by SCART and the effect size for our precondition results indicating that phrases were encoded as chunks ($\eta^2 = .436$ for recall of critical words in chunkable vs. unchunkable lists). The null effects of list-type strongly contradict the freed-capacity prediction that chunking serves to free processing capacity for encoding additional words in rapidly presented lists (especially for young adults): Contrary to SCART, no more extraneous words were recalled in phrase than no-phrase lists for either young or older adults. Because recoding of words into phrases did occur in the phrase lists (see the precondition results), this indicates that storage and processing capacity was not freed up when smaller chunks were recoded into larger chunks in Experiment 1.

The all-chunks-are-equal assumption. For both young and older adults, recall was better for the two-word super-chunks in Phrase-lists than for isolated words in corresponding list-positions in No-Phrase lists t (60) = 10.584, p <.001, and this finding also held, t (60) = 7.002, p <.001, for the identical critical words in Phrase and No-Phrase lists (see Table 4), indicating that the superior recall of words in chunks was not due to the (single) words that differed in two- versus one-word chunks. These data are difficult to explain under the all-chunks-are-equal assumption, which predicted no difference in recall of *night* in *night gown* versus as an isolated word in No-phrase lists. However, these data further support the basic premise underlying the freed-capacity assumption in CART: Because recall was better for two-word than one-word chunks, even greater memory capacity should have become freed up for recalling extraneous words in Phrase than No-phrase lists, making the present (null) results all the more surprising.

List-length results for extraneous word recall. Figure 2 (left panel) shows the list-length predictions of SCART for extraneous word recall by capacity-3, -4, and -5.5 participants: an effect of age; an effect of list-length, with more extraneous words recalled in 7- than 6-word lists; and better recall of extraneous words in longer than shorter lists, especially for young adults. Figure 2 (right panel) shows the actual mean recall of extraneous words by age and list-length. A 2 (age group: young vs. older) x 2 (list length: 6 vs. 7 words) ANOVA on these data yielded a main effect of list length, F (1,59) = 32.539, MSE = .071, p < .001, with more extraneous words recalled in 7- than 6-word lists, but no main effect of age, F (1,59) = 1.254, MSE = .188, p = .267, and no age x list length interaction, F < 1. The absence of an age x list length and the null effect of age on recall of extraneous words are inconsistent with SCART, but the main effect of list length comports with SCART prediction. However, a simpler explanation of this list length effect is that longer lists contained more extraneous words and therefore offered more opportunities for recall of extraneous words.

Similar list-length results were obtained in our analysis that included high capacity young adults as a separate group to test the interaction between list length and capacity indicated by SCART. A 3 (capacity: low (older adults) vs. medium (below-median young adults) vs. high (above-median young adults) x 2 (list length: 6 vs. 7 words) ANOVA found a main effect of list length, F (1,58) = 31.246, MSE = .070, p < .001, with more extraneous words recalled in 7- than 6-word lists, a main effect of capacity F (2,58) = 8.149, MSE = .153, p = .001, but no capacity x length interaction, F (2,58) = 1.062, MSE = .070, p = .352. The absence of a capacity x list length interaction on recall of extraneous words for high-capacity young adults contradicts the list-length prediction of SCART that recall of extraneous words will increase with list length, especially for young adults with high capacity.

List-length results for recall in chunks. Figure 3 (left panel) shows the list-length predictions of SCART for chunk recall by capacity-3, -4, and -5.5 participants: a main effect of age; and more chunks recalled (with phrases in phrase list counted as single chunks) as a function of list-length for young adults (capacity-4 or greater), but not older adults. Figure 3 (right panel) shows the data for mean chunks recalled by age and list-length. A 2 (age group: young vs. older) x 2 (list type: phrase vs. No-phrase) x 2 (list length: 6 vs. 7 words) ANOVA on the number of chunks recalled yielded a main effect of length, F (1, 59) = 7.789, MSE = .211, p = .007, with more chunks recalled in 6- than 7-word lists, but no main effect of age, F (1, 59) = 1.113, MSE = 1.488, p = .296, and no age x list-length interaction, F < 1. The null effect of age contradicted the Resource-capacity prediction that capacity measured in chunks is significantly smaller for older than young adults.

To test the large age x list-length interaction predicted under SCART for high-capacity young adults (see capacity-5.5 in Figure 3, left panel), we next analyzed these results with capacity (low vs. medium vs. high) as a between-subjects factor. A 3 (capacity: low (older adults) vs. medium (below-median young adults) vs. high (above-median young adults) x 2 (list length: 6 vs. 7 words) ANOVA found a main effect of list length, F (1,58) = 8.326, MSE = .111, p = .005, with more chunks recalled in 6- than 7-word lists, the expected main effect of capacity F (2,58) = 6.764, MSE = .507, p < .001, but no capacity x length interaction, F < 1. This ANOVA replicated the finding of the two-group ANOVA reported in the previous paragraph; both the absence of a list length x capacity interaction when high-capacity young adults were included in the analysis as a separate group and the decrease in chunk-recall with increasing list length for high-capacity young adults (6-word lists: M = 3.805, SD = .449; 7-word lists, M = 3.640, SD = .475) contradict the predictions of SCART.

Of course, SCART could be modified to predict not fixed recall in chunks but a fixed recall *proportion* in chunks presented for lists of a given length, in which case effects of list length and list type on recall may reflect differences in the number of chunks presented and potentially recallable. For example, length-6 phrase lists contain 4 chunks, but length-6 no-phrase lists contain 6 chunks that are potentially recallable. We therefore analyzed the number of chunks recalled as a proportion of presented chunks, with mean proportions shown in Table 3. A 2 (age group: young vs. older) x 2 (list type: phrase vs. No-phrase) x 2 (list length: 6 vs. 7 words) ANOVA on these data yielded the same main effect of length, F (1, 59) = 138.500, MSE = .008, p < .001, but no effect of age, F < 1, and no reliable interactions. The main effect of length again contradicts the general assumption of SCART that capacity measured at supra-span levels is independent of stimulus factors such as list-length.

List-type results for recall in chunks. Figure 4 (left panel) shows the list-type predictions of SCART for chunk recall for capacity-3, -4, and -5.5 participants: a main effect of age; and more chunks recalled in no-phrase than phrase lists for young adults (capacity-4 and greater), but not older adults. Figure 4 (right panel) shows the actual mean number of chunks recalled by age and list-type. The 2 (age) x 2 (list type) x 2 (list length) ANOVA on chunk recall (described earlier) yielded a main effect of list type, F (1, 59) = 381.800, MSE = .200, p < .001, with fewer chunks recalled in phrase than no-phrase lists. This finding contradicts the chunk-constancy for list-type prediction of SCART.

In addition this ANOVA yielded two significant interactions involving list type: One was an age x list type interaction F (1, 59) = 4.092, MSE = .200, p = .048, which reflected a smaller mean difference between list types for older adults (M = .988, SD = .467) than young adults (M = 1.208, SD = .359); the other significant interaction in the ANOVA involved list length x list type, F (1, 59) = 18.129, MSE = .190, p < .001, reflecting a larger difference in recall between phrase versus no-phrase lists for list-length 6 than 7. The theoretical implications of these interactions are discussed in MacKay, Hadley, and Abrams (submitted).

Presentation-rate results for recall in chunks. To test the fixed-capacity assumption that chunk capacity is independent of presentation rate, we computed a 2 (age) x 2 (list-type) x 2 (list-length) x 2 (rate: fast versus slow) ANOVA on Experiment 1 data with 170 and 200 ms/word corresponding to the fast and slow rates for young adults and 200 and 270 ms/word corresponding to the fast and slow rates for older adults. In addition to the effects described earlier, this ANOVA yielded a main effect of rate, F (1, 59) = 6.618, MSE = .358, p = .013, which contradicts the assumption that capacity is independent of presentation rate factors.

Precondition Results supporting the Resource-capacity Analyses

The ceiling precondition. Experiment 1 data met the ceiling precondition for testing SCART since recall of extraneous words was 57% overall, and well below ceiling for both young adults (61%) and older adults (54%). This indicates that the lists exceeded memory span for all participants, enabling possible benefits in recall due to freed-capacity.

The chunking precondition. Six sources of data indicated that Experiment 1 data met the chunking precondition. First, results from two conditional recall analyses supported the chunking precondition. If A represents either the first or second word in a chunk in phrase lists and B represents the other word, the first conditional analysis indicated that given recall of A in a chunk in phrase lists, B was recalled with higher probability than if A was missed for both young adults, t (27) = 6.532, p < .001 and older adults, t (29) = 9.426, p < .001. The second conditional analysis indicated that given recall of A, B was recalled with significantly higher probability than the word in the same list-position in the corresponding no-phrase list for both young adults, t (29) = 7.123, p < .001 and older adults, t (30) = 9.426, p < .001 (for the preceding analyses, the degrees of freedom varied somewhat due to patterns of empty cells in the data sets where participants did not have any responses for a given conditional analysis).

Also supporting the chunking precondition, more words were recalled in phrase than no-phrase lists, t (60) = 11.466, $p < .001$, as would be expected if the two-word chunks were acting as single units. Fourth, as already noted, critical words were recalled better in two-word chunks than in identical list-positions in No-Phrase lists for both young and older adults, outcomes that could only occur if participants had encoded chunkable words as super-chunks rather than independent words. Fifth, two-word chunks were better recalled than one-word chunks even when we gave one-word chunks words a primacy advantage by excluding from analysis list-final one-word chunks t (60) = 10.774, $p < .001$. Sixth, serial order of recall indicated many word-order reversals for isolated words, but virtually none inside two-word chunks, which again suggests recoding into familiar two-word phrases (see also Miller, 1958).

The perception precondition. Support for the chunking precondition also supports the perception precondition for testing SCART: For example, if words were misperceived with high probability in Experiment 1, then critical words would be encoded and recalled no better in phrase than no-phrase lists since chunking is a post-perceptual effect: Only if the critical words in two-word chunks were first perceived as words could they be encoded and recalled as phrases containing those words. However, our results indicated that critical words were *better* recalled in phrases than as isolated words, suggesting that critical words were not misperceived with high probability in either phrase or no-phrase lists.

Other sources of evidence suggested that presentation rates in Experiment 1 (170-270 ms/word) were sufficiently slow that perceiving the words was not a problem. For example, phonologically similar word-substitutions, a standard index of perceptual difficulties, were equally frequent in Experiment 1 and 2 (approximately 8% of all trials), a finding that rules out perceptual difficulty as an issue because the presentation rate in Experiment 2 was so slow (2000ms/word) as to make misperception a non-issue.

The strategy precondition. We found no evidence for age-linked strategies that might have affected recall of extraneous words and contributed to the failure of the freed-capacity prediction for young or older adults. To assess strategy differences, we analyzed recall as a function of input position: initial-position (the first word in the lists), final (the last word in the lists), and middle-position. At least two other words always preceded and followed extraneous words in middle-positions, which varied with chunk locus in the same way for both phrase and no-phrase lists. Depending on chunk locus, middle-positions included words 3 or 4 in 6-word lists and words 3, 4, or 5 in 7-word lists. Table 5 shows mean recall of extraneous words (in %) as a function of age for three list positions in phrase vs. No-phrase lists in Experiment 1. A 2 (age: young vs. older) x 2 (list type: phrase vs. No-phrase) x 3 (list position: initial, middle, final) mixed ANOVA on these data yielded the same effects as the 2 x 2 ANOVA described earlier, but in addition, a main effect of list position, F (2, 118) = 82.893, MSE = .084, p < .001, with better recall for initial than middle or final words, F (1, 59) = 155.622, MSE = .109, p < .001, which did not differ from each other in Helmert contrasts, F <1. In addition, there were no interactions between position and age, F (2, 118) = 2.036, MSE = .087, p = .135, or list type F (2, 118) = 1.924, MSE = .041, p = .151, and the three way interaction between position, age and list type was not significant, $F < 1$. The lack of reliable interactions involving list position was consistent with the hypothesis of Murdock (1974, p. 168) that rehearsal strategies are not possible with presentation rates of 200/ms/word or less. Present results therefore do not support the hypothesis that age-linked strategies for recall or rehearsal contributed to the failure of the freed-capacity prediction in Experiment 1.

Experiment 2: Slow Presentation Rate

Experiment 2 had the same goals as Experiment 1, plus several new ones. One was to replicate with better procedures the main finding in previous studies that supported the fixed-capacity assumption. Tulving & Patterson (1968, p. 239) express the finding succinctly: "The number of retrieved functional units is independent of the size of the units" in immediate recall tasks (see also Bower, 1972; and Tulving & Pearlstone, 1966). Because presentation rates were approximately 2.0 sec/chunk in these earlier studies, Experiment 2 adopted 2.0 sec/word as the presentation rate.

Experiment 2 also tested whether previous results violated the strategy precondition for valid tests of the fixed-capacity and freed-capacity assumptions. We recognized that participants may adopt rehearsal strategies with presentation rates of 2.0 sec/word, and we were especially interested in a possible "chunk-specific” strategy that could in principle yield artifactual support for the freed-capacity prediction. Participants adopting this strategy spend less time rehearsing the familiar phrases in phrase lists and more time rehearsing the remaining extraneous words because they realize that two-word chunks are easier to recall than isolated words (consistent with Experiment 1 results). These participants will therefore recall more extraneous words in phrase than no-phrase lists, not due to freed-capacity, but because they have selectively rehearsed extraneous words in phrase but not no-phrase lists. To assess this and other possible strategies, we queried Experiment 2 participants about strategy use in post-experimental questionnaires. Other procedures were identical to Experiment 1, updating earlier procedures where participants turned over cards in a prepared deck (Tulving & Patterson, 1968; and Tulving & Pearlstone, 1966), or saw cards containing 4 letter strings for about 5 seconds apiece (Bower, 1972).

Method

Participants

Participants were 32 UCLA undergraduates who received course credit for participation (mean age = 19.5 years, SD= 1.82 years) and 34 healthy, community-dwelling older participants (mean age = 73.7 years, SD = 4.50 years) from the UCLA Cognition and Aging Laboratory participant pool who received $10.00/hour for participating. All participants reported normal or corrected-to-normal vision and were fluent speakers of English. Table 1 contains education, health, digit span, and Nelson-Denney vocabulary information for these groups.

Materials and Design

Materials were adaptations of the 32 lists in Experiment 1 that accomplished three goals: to make the chunked and extraneous words identical in Experiments 1 and 2, to avoid possible ceiling effects at the slower presentation rate in Experiment 2, and to vary list length over a wider range than in Experiment 1. To achieve these goals, Experiment 2 lists were identical to Experiment 1 lists except that 1-4 new unrelated words within the same frequency range were added to the ends of the original lists, resulting in 32 lists evenly divided between 7, 8, 9, and 10 words in length with no critical or phrase words beyond word 7 (see Appendix).

Procedure

Each participant saw 16 phrase lists and 16 no-phrase lists with RSVP procedures identical to Experiment 1 except that words appeared at 2000 ms/word. In addition, participants were not instructed regarding phrases or possible strategies for aiding recall.

Results and Discussion

Experiment 2 data met the perception, ceiling and chunking preconditions for testing SCART, but not the strategy precondition. However, we again discuss our main results before precondition results to simplify exposition since the strategy analyses did not change the interpretation of the main results.

Main Results

Table 2 shows the mean number of chunks correctly recalled by age group and list type for all lists in Experiment 2. Significantly fewer chunks were recalled in phrase than no-phrase lists for both young adults, t (31) = -9.435, $p < .001$, and older adults, t (33) = -4.683, $p < .001$, and the overall mean chunks recalled in no-phrase lists was about 4 chunks for older adults and about 5 chunks for young adults. We therefore initially compared the data with capacity-4 predictions of SCART for older adults and with capacity-5 predictions for young adults.

Because SCART predicted measurably larger interactions in our data for capacity-6 or greater (see Figures 7-10), we subsequently analyzed separately the data for young participants with mean recall of 6 chunks or greater to determine whether these interactions emerged with larger capacities. However, only 3 young participants had mean recall of 6 chunks or greater, we performed a median split of the data for young participants based on mean capacity per participant (determined by performance for the no-phrase lists) and conducted additional analyses with capacity as a between-subjects factor. As in Experiment 1, these capacity analyses yielded no new main effects or interactions of interest; for the sake of clarity, they are not reported here.

Freed-capacity results. Figure 5 shows the mean recall of extraneous words by age and list-type (right panel), together with predicted recall in SCART for capacity-3.5, -4.5, and -5.5 participants (left panel): an effect of age with better recall of extraneous words for young than older adults; and better recall of extraneous words in phrase than no-phrase lists, especially for young adults. A 2 (age: young vs. older) x 2 (list type: phrase vs. No-phrase) mixed ANOVA on these data yielded a main effect of age, F (1, 61) = 15.054, MSE = .0261, p < .001, with better recall of extraneous words for young than older adults, and a marginal age x list type interaction, F (1, 61) = 2.842, MSE = .0514, p = .097 (see Figure 7), but no main effect of list type, F (1, 61) = .020, MSE = .0514, p = .887. The null effect of list type was not due to insufficient power in our design. Power analysis indicated a power of .883 to detect a large effect as predicted by SCART and the effect sizes for our precondition results indicating that phrases were encoded as chunks (η^2 = .692 for recall of critical words in chunkable vs. unchunkable lists).

These list-type results contradict the freed-capacity assumption that recoding into larger chunks serves to free capacity for recalling additional single-word chunks, especially for young adults: Because phrase words were recoded into phrases in phrase lists (see the precondition results following this section), the fact that no more extraneous words were recalled in phrase than no-phrase lists for either young or older adults indicates that storage and processing capacity were not freed up when participants recoded smaller chunks into larger chunks. The marginal age x list-type interaction did not qualify this conclusion since SCART predicted an interaction in the opposite direction (see Figure 5).

The all-chunks-are-equal assumption. The all-chunks-are-equal assumption in CART predicted no recall difference between one- and two-word chunks in identical list-positions. Table 4 shows the mean probability of recall for the four words in the two-word chunks of phrase lists and the corresponding four words in the same list positions in no-phrase lists for young and older adults. For both young and older adults, recall was better for words in two-word chunks than in identical list-positions in No-Phrase lists, t (65) = 6.848, $p < .001$. This superior recall of two-word chunks was not due to lexical-level differences between two- versus one-word chunks since the identical (critical) word was recalled better as part of a two-word chunk in Phrase lists than as an isolated word in No-Phrase lists t (65) = 11.828, $p < .001$. These results are difficult to explain under the all-chunks-are-equal assumption but further support the basic premise underlying predictions based on the freed-capacity assumption in SCART: Because recall was better for two-word than one-word chunks, even more memory capacity should have been freed up for storing extraneous words in Phrase than No-Phrase lists, making the present (null) results all the more surprising.

List-length results for extraneous words. Figure 6 (left panel) shows the list-length predictions of SCART for recall of extraneous words: an effect of age, with reduced recall of extraneous words for older adults; and better recall of extraneous words in longer than shorter lists, especially for young adults. Figure 6 (right panel) shows the actual mean recall of extraneous words by age and list-type. A 2 (age group: young vs. older) x 4 (list length: 7-10 words) ANOVA on these data yielded main effect of age, $F(1, 64) = 11.47$, $MSE = 1.058$, $p = .001$, with older adults recalling fewer extraneous words overall, and a main effect of length, $F(3, 192) = 11.47$, $MSE = .255$, $p < .001$, with more extraneous words recalled in longer lists. There was no age x list length interaction, F < 1, and the stronger age x list length interaction predicted under SCART for capacity-6 young adults did not emerge in a separate that included high-capacity young adults as a separate group. As in Experiment 1, the effect of length on recall of extraneous words does comport with SCART predictions, but may also be an effect of the increasing number of extraneous words available in the longer lists.

List-length results for recall in chunks. Figure 7 (left panel) shows the list-length predictions of SCART for the capacity-3.5, -4.5, and 5.5 participants: a main effect of age; and better chunk-recall with list-length 7 than 8 for young adults, but no effect of list-length for older adults. Figure 7 (right panel) shows the actual mean number of chunks recalled by age and list-length. To test the chunk-constancy prediction for list length, conducted a 2 (age group: young vs. older) x 4 (list length: 7, 8, 9, 10 words) ANOVA. This analysis indicated a main effect of age $F(1, 65) = 36.301$, $MSE = 2.092$, $p < .001$, and a marginal list length x age interaction, $F(3, 195) = 2.281$, $MSE = .159$, $p = .081$, but no main effect of length, $F < 1$. The marginal age x list length interaction appears to reflect a linear increase in the difference between the number of chunks recalled by young and older adults as a function of list length, $F(1, 65) = 3.784$, $MSE = .197$, $p = .056$, a trend that was not predicted by SCART (see Figure 7, left panel).

List-type results for recall in chunks. Figure 8 (left panel) shows the predictions for list-type under SCART for capacity-3.5, -4.5, and 5.5 particpants: a main effect of age; and better chunk recall in no-phrase than phrase lists, but only for young adults. Figure 8 (right panel) shows the actual mean number of chunks recalled by age and list-type. A 2 (list type) x 2 (age group) ANOVA on the data in Figure 10 yielded a main effect of list type, F (1, 65) = 81.898, MSE = .0977, p < .001, with fewer chunks recalled in phrase than no-phrase lists, and a main effect of age, F(1, 65) = 36.298, MSE = 1.046, p < .001, with fewer chunks recalled by older than young adults, and an age x list type interaction, F (1, 65) = 5.346, MSE = .0952, p = .024, which reflected a smaller mean difference in recall between list types for older (M = .3640, SD = .4837) than younger adults (M = .6138, SD = .3946). The main effect of age on number of chunks recalled comports with SCART, but not the main effect of list type or the age x list type interaction, since SCART assumes that capacity measured in chunks is fixed and independent of stimulus conditions.

Precondition Results supporting Resource-capacity Analyses

The ceiling precondition. Experiment 2 data met the ceiling precondition for testing SCART since recall of extraneous words was 56% overall, and well below ceiling for both young adults (61%) and older adults (51%). This indicates that the lists exceeded memory span for all participants, enabling possible benefits to recall under the freed-capacity assumption.

The chunking precondition. Six sources of data indicated that Experiment 2 data met the chunking precondition. First, if A is either the first or second word in a two-word chunk and B is the other word, conditional analyses indicated that given recall of A, B was recalled with significantly higher probability than conditional recall of the word in the same list-position in the corresponding no-phrase list (with list-initial words excluded to minimize primacy effects), t (65) = 12.332, $p < .001$. Second, given recall of A in a two-word chunk, B was recalled with higher probability than if A was missed, t (65) = 13.967, $p < .001$. Third, participants recalled more words in phrase than no-phrase lists, t (65) = 14.575, p < .001, as should occur when two-word chunks act as single units. Also, for both young and older adults, words were better recalled in two-word chunks than in the corresponding list-positions in no-phrase lists, t (60) = 6.848, p < .001, which could only occur if participants were encoding the two-word chunks as single units rather than two independent words. Finally, the superior recall of 2-word chunks was not due to the differing words in two-word versus one-word chunks since the identical (critical) word was recalled better as part of a two-word chunk in phrase lists than as a single word in no-phrase lists, t (65) = 11.828, p < .001. These results indicate that two-word chunks were better recalled than one-word chunks, a finding that also held true for one- versus two-words chunks within phrase lists, despite a primacy advantage for extraneous words (see the primacy effect discussed in the next section). To achieve this primacy advantage, we excluded all extraneous words in list-positions 7-10 and compared recall for the remaining extraneous words in phrase lists (Mean list position = 3.34) with the four words in chunks (Mean list position = 3.69). The words in chunks were recalled better than the extraneous words, t (65) = 11.966, $p < .001$, a difference that is difficult to explain in SCART, where a chunk is a chunk, regardless of "size."

The perception precondition. Experiment 2 data satisfied the perception condition because word perception is not an issue at 2000ms/word, indicating that word misperception cannot explain why the freed-capacity prediction failed.

The strategy precondition. Experiment 2 results failed to satisfy the strategy precondition for list-lengths 8-10. As in Experiment 1, we examined recall as a function of list position for young versus older adults to test for possible age-linked strategies. Initial and final list positions were defined as in Experiment 1, but middle positions varied with list length: positions 3-5 for length-7 lists, positions 4-5 for length-8 lists, positions 4-6 for length-9 lists, and positions 4-7 for length-10 lists. A 2 (age: young vs. older) x 2 (list type: phrase vs. No-phrase) x 3 (list position: initial, middle, final) ANOVA on these data yielded the same effects as the earlier age x list type ANOVA, but in addition, a main effect of list position, $F(2, 128) = 57.444$, $MSE = .0665$, $p < .001$, with better recall for initial than middle and final positions, and an age x list position interaction, $F(2, 128) = 3.276$, $MSE = .0665$, $p = .041$, reflecting much better recall of list-middle words for young than older adults, $F(1, 64) = 10.618$, $MSE = .0164$, $p = .002$, a smaller age difference for list-initial words, $F(1,64) = 3.912$, $MSE = .0649$, $p = .05$, and no age difference for list-final words, $F<1$. Because this age x list position interaction suggested strategy effects, we analyzed our post-experimental questionnaires for verification. For participants who reported strategy use, 43 % claimed to focus on recall of list-final words or list-initial words or both, strategies likely to reduce recall of list-middle words, especially for longer lists.

To evaluate effects of such strategies on our freed-capacity results, we examined recall of extraneous words by age and list position (initial, middle, and final) in short 7-

word lists versus longer 8-10 word lists. The length 8-10 lists exhibited the age x list position interaction of the earlier age x analysis but not the length 7 lists. To rule out strategy effects, we therefore repeated our earlier freed-capacity analysis using only the data from the 7-word lists, which also met the other preconditions for testing the freed-capacity prediction: Critical words were recalled better in phrase than no-phrase lists, t (65) = 7.741, $p < .001$, and participants recalled 55% of the extraneous words in 7-word lists, sufficiently below ceiling to allow chunking-related improvement due to freed capacity.

A 2 (age: young vs. older) x 2 (list type: phrase vs. No-phrase) mixed ANOVA on data for extraneous words in length-7 lists again yielded a main effect of age and a marginal age x list type interaction, F (1, 61) = 3.525, MSE = .0258, p= .065, reflecting slightly poorer recall of extraneous words in phrase lists for older adults. However, there was no main effect of list type, F (1, 61) = .006, MSE = .0258, a result that replicates Experiment 1 and Experiment 2 results. This result contradicts the hypothesis that chunking serves to free resources within a fixed-capacity store, and renders unlikely the possibility that the freed-capacity prediction failed because of strategy effects.

Results Compared across Experiments 1 and 2

The integral-chunk results

To test the integral-chunk assumption that chunks are unitary and that storage capacity is an integral number of chunks, we examined the frequency distribution of individual capacity limits, defined as the mean recall in chunks for Unchunked lists, which fell below ceiling for all participants in Experiments 1 and 2. The distribution patterns for young and older adults were similar and are combined in Figure 9, which shows the distributions for participants with integer capacities, where capacities within + or - .167 chunks of an integer were counted as falling on the integer, and with off-integer capacities, i.e., all other chunk capacities, which Figure 9 represents as falling at 2.5, 3.5, 4.5, 5.5, or 6.5 chunks. Fewer participants had integer capacities than off-integer capacities in both Experiment 1 and 2, t (126) = -16.126, $p < .001$, with no indication of multimodal peaks at the integer values in either frequency distribution, outcomes contrary to the integral-chunk assumption.

Effects of Rate on Recall in Chunks

To test the assumption that chunk capacity is independent of the differing presentation rates in Experiment 1 versus 2, we compared mean correct recall in chunks for the identical length-7 lists in Experiment 1 (fast presentation) versus 2 (slow presentation). A 2 (age) x 2(rate: fast vs. slow) x 2 (list-type: phrase vs. No-phrase) ANOVA on these data yielded no reliable list-type x rate or age x list-type x rate interactions, but an effect of rate F (1, 124) = 45.824, MSE = 1.074, $p < .001$, with more chunks recalled at the slow than fast rate, and an age x rate interaction, F (1, 124) = 11.362, MSE = 1.074, $p < .001$, such that the difference in recall for young vs. older adults was larger at the slow than at the fast rate. Neither the main effect of rate nor the age x rate interaction comport with the assumption that recall in chunks is independent of presentation rate.

Effects of Rate on the Age x List-type Interaction

Compared to young adults, older adults exhibited smaller effects of list type on the number of chunks recalled in both Experiment 1 and 2. To compare the strength and detailed nature of these age x chunk type interactions across these experiments, Figure 10 includes the data only for critical words (which were identical in phrase and no-phrase lists), and only for length-7 lists (which met the strategy precondition in both Experiment 1 and 2), and excludes the data for critical words in initial and final positions in the lists (to eliminate primacy and recency effects). Despite the reduced power, a 2 (age: young vs. older) x 2 [rate: Experiment 1 (fast) versus Experiment 2 (slow)] x 2 (list type: Phrase vs. No-phrase) ANOVA comparing recall of critical words for 7-word lists in both experiments (see means in Figure 10) yielded the usual main effect of age, $F(1, 123) = 13.338$, $MSE = .0469$, $p < .001$, a main effect of rate, $F(1, 123) = 76.709$, $MSE = .0469$, $p < .001$, with better recall at the slow than fast rate, and a main effect of list type, $F(1, 123) = 45.040$, $MSE = .0316$, $p < .001$, with better recall in Phrase than No-phrase lists, plus exactly two significant interactions. One was the age x list type interaction seen in the earlier age x list type analyses, $F(1, 123) = 6.293$, $MSE = .0316$, $p = .013$, reflecting relatively better recall by older adults of critical words in Phrase than No-phrase lists. Also reliable was the list type x rate interaction, $F(1, 123) = 8.255$, $MSE = .0316$, $p = .005$, reflecting relatively better recall of critical words in Phrase than No-phrase lists at the slower rate.

Experiment 3: Alternate Hypotheses: Phonological Loop versus Binding

Why did participants recall more words in Phrase than No-phrase lists in Experiments 1 and 2? Experiment 3 tested alternate hypotheses derived from two theoretical conceptions of storage capacity that contrast both with each other and with CART. One was working memory theory, where a phonological loop stores phonological information for a fixed period of time (about 2 seconds, e.g., Baddeley, 1986, Baddeley & Hitch, 1975). Because the phonological loop simply maintains phonological information without encoding syntactic or semantic information under working memory theory, rate of articulation determines how much information can enter the phonological loop for immediate recall. By hypothesis, then, participants recalled more words from Phrase than No-phrase lists because articulation rates were faster for Phrase than No-phrase lists in Experiments 1-2.

To test this phonological loop hypothesis, we recorded young and older participants reading the experimental lists in Experiment 2 as rapidly as possible. We then digitized their outputs to determine articulation rates for Phrase versus No-phrase lists. The phonological loop hypothesis predicted reliably faster articulation rates for Phrase than No-phrase lists.

The binding hypothesis provides a contrasting theoretical conception. Unlike the phonological loop hypothesis, semantic-level units play a critical role in recalling unrelated words and phrases in immediate memory tasks under the Binding hypothesis. That is, semantic-level units must become bound to their list-context within the available encoding time in immediate memory tasks because the cue for retrieval is list-context and only those semantic-level units with sufficiently strong links to their list-context will be recalled. Native speakers normally acquire the semantic-level unit for a familiar word or phrase as a child and list-context refers to the occurrence of that semantic-level unit in a particular episodic context or list. Thus, more words were recalled in Phrase than No-Phrase lists under the binding hypothesis because one link between the list-context and a semantic-level unit was sufficient to recall each two-word phrase in Phrase lists, whereas two links between list-context and semantic-level units were necessary to recall the corresponding isolated words in No-Phrase lists. No relation between articulation rate and word recall in Phrase versus No-phrase lists is therefore expected under the binding hypothesis, unlike the phonological loop hypothesis.

Method

Participants

There were 19 participants: 12 UCLA undergraduates and 7 older adults. None had participated in Experiments 1 or 2. All were native speakers of English, reported normal or corrected-to-normal vision, and received either $10 or class credit for participating.

Materials

The materials were the 32 experimental lists in Experiment 2.

Procedure

Each participant saw 16 Phrase and 16 No-phrase lists in 36-point font presented in random order via Macintosh Computer running Psyscope. The instructions resembled Experiment 2 except that participants were asked to read the lists aloud as quickly as possible without making errors. Prior to each list, the warning “Get ready for the next list!” appeared for two seconds, followed by a 1200 ms pause, then the list. To facilitate rapid articulation, the words appeared in a vertical column down the center of the screen. After reading a list, participants pressed the space bar to trigger the next list. As in Experiments 1 and 2, four practice lists and reminder instructions preceded the first experimental list.

Results

We only scored data from error free lists to maximize our ability to observe rate differences between Phrase versus No-phrase lists. After digitizing each error-free list using Soundedit 16, we measured list duration and computed articulation rate in syllables/second by dividing list duration for each participant by the total number of syllables in the list. We then computed mean articulation rate by list type and age group with the results shown in Figure 12.

Mean articulation rate for Phrase lists (M = 3.08 syllables/sec, SD = .747) and No-phrase lists (M = 3.04 syllables/sec, SD = .748) was nearly identical, with the mean within-subject difference in articulation rate negligibly faster (0.0386 syllables/sec) for Phrase than No-phrase lists. A 2 (age group: young vs. older) X 2 (list type: Phrase vs. No-phrase) mixed ANOVA on the mean articulation rates yielded no effect of list type, $F(1, 17) = .293$, $MSE = .0965$, $p = .595$, and no age x list type interaction, $F(1, 17) = .423$, $MSE = .0965$, $p = .524$, but articulation rates were reliably slower for older than young adults, $F(1, 17) = 46.350$, $MSE = .291$, $p < .001$.

We next estimated the size of the phonological loop in seconds for young and older participants in Experiment 2 by dividing the mean number of syllables recalled for list types in Experiment 2 by the mean articulation rate for that list type and age group in Experiment 3. Figure 12 shows the mean estimated phonological loop sizes as a function of list type and age group. A 2 x 2 ANOVA on these data indicated reliably larger estimated phonological loop sizes for Phrase than No-phrase lists, $F(1, 17) = 10.348$, $MSE = .247$, $p = .005$, and for older than young adults, $F(1, 17) = 26.703$, $MSE = .0567$, $p < .001$, with no reliable age group x list type interaction, $F(1, 17) = 1.391$, $MSE = .247$, $p = .224$.

References

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

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

Bower, G.H. (1972). Perceptual groups as encoding units in immediate memory. Psychonomic Science, 27, 217-219.

Cohen, J.D., MacWhinney, B., Flatt, M., & Provost J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25(2),* 257-271.

Kucera, H., & Francis, W.N. (1967). *Computational analysis of present-day American English.* Brown University Press, Providence, RI.

MacKay, D.G., Hadley, C.B., & Abrams, L. (submitted). Beyond soup stones: Testing the core assumptions of age-linked resource-capacity theories.

Murdock, B.B. (1974). *Human memory: Theory and data.* Oxford, England: Lawrence Erlbaum.

Tulving, E. & Patterson, R.D. (1968). Functional units and retrieval processes in free recall. *Journal of Experimental Psychology*, *77*, 239-248.

Tulving, E. & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, *5*, 381-391.

Table 1

Background Information for Participants in Experiments 1-2, with standard deviations (SD) in parentheses

	Experiment 1		Experiment 2	
	Young Adults N=30	Older Adults N=31	Young Adults N=32	Older Adults N=34
Mean Age	19.4 (2.32)	70.5* (4.46)	19.5 (1.82)	73.68* (4.50)
Mean years of education	13.5	14.6	13.8	15.8
Mean Health Self-rating[1]	7.9	8.1	7.9	7.3
Mean Forward Digit-span[2]	7.7	6.2*	7.4	7.3
Mean Backward Digit-span[3]	5.5	4.3*	5.5	5.7
Mean Nelson-Denney[4]	16.25	20.1*	14.5	22*

*Differences between age groups significant at p = .05 or less

[1] Self-rate health relative to "a person of your age" on a 10-point scale with "1 = poor" and "10 =excellent."

[2] Maximum number of auditorily presented digits recalled correctly in the same order as presented.

[3] Maximum number of auditorily presented digits recalled correctly in the reverse order as presented.

[4] Vocabulary skills test with maximum score 25.

Table 2

Mean number of chunks correctly recalled by age group and list type for all lists and 7-word lists in Experiments 1 and 2. SDs are in parentheses.

		Number correct Experiment 1		Number correct Experiment 2	
Age group	List type	Length-7	All Lists	Length-7	All Lists
Young Adults	Phrase	2.72 (.559)	2.69 (.560)	4.02 (.569)	4.28 (.687)
	No-phrase	3.66 (.787)	3.90 (.682)	4.97 (.888)	4.89 (.795)
Older Adults	Phrase	2.66 (.601)	2.62 (.528)	3.27 (.762)	3.31 (.637)
	No-phrase	3.35 (.804)	3.605 (.701)	3.83 (.680)	3.706 (.932)

Table 3

Mean proportion of chunks recalled (with Phrases scored as single chunks) by list type and age group in Experiment 1 (fast presentation) and Experiment 2 (slow presentation). SDs are in parentheses.

		Experiment 1	Experiment 2
Age group	List type	Proportion correct	Proportion correct
Young adults	Phrase	.704 (.106)	.670 (.104)
	No-phrase	.601 (.103)	.583 (.092)
Older Adults	Phrase	.677 (.101)	.525 (.100)
	No-phrase	.552 (.106)	.441 (.111)

Table 4

Mean probability of recall for critical words and both words in the two-word chunks of Phrase lists and the corresponding words in the same list positions in no-Phrase lists by age and list type in Experiment 1 (fast presentation, all lists) and Experiment 2 (slow presentation, all lists). Standard deviations (SDs) are in parentheses

		Experiment 1		Experiment 2	
Age group	List type	Proportion correct		Proportion correct	
		Critical words	Both words	Critical words	Both words
Young adults	Phrase	.712 (.127)	.652 (.159)	.815 (.116)	.722 (.124)
	No-phrase	.562 (.167)	.421 (.158)	.558 (.126)	.464 (.147)
Older Adults	Phrase	.691 (.144)	.629 (.158)	.688 (.102)	.579 (.128)
	No-phrase	.544 (.155)	.387 (.147)	.484 (.126)	.300 (.146)

Table 6

Mean probability of recall for extraneous words as a function of list position (initial, middle, and final) for young versus older adults in Experiment 1 (fast presentation, all lists) and in Experiment 2 (slow presentation, all lists, length 7 lists, and lengths 8-10 lists). SDs are in parentheses

			Experiment 1	Experiment 2		
Age group	List type	List Position	Correct Recall Probability	Correct Recall Probability		
			All List-lengths	All List-lengths	List-length 7	List-lengths 8-10
Young Adults	Phrase	Initial	.934 (.214)	.830 (.243)	.922 (.345)	.781 (.282)
		Middle	.414 (.257)	.559 (.175)	.609 (.356)	.517 (.187)
		Final	.462 (.235)	.418 (.209)	.625 (.311)	.424 (.232)
	No-Phrase	Initial	.839 (.230)	.800 (.238)	.891 (.230)	.76 (.277)
		Middle	.486 (.213)	.493 (.175)	.578 (.362)	.47 (.198)
		Final	.521 (.236)	.517 (.221)	.572 (.271)	.502 (.238)
Older Adults	Phrase	Initial	.775 (.213)	.689 (.239)	.757 (.339)	.624 (.274)
		Middle	.476 (.258)	.359 (.175)	.367 (.349)	.329 (.181)
		Final	397 (.235)	.430 (.204)	.44 (.310)	.388 (.227)
	No-Phrase	Initial	.739 (.213)	.693 (.239)	.829 (.297)	.629 (.274)
		Middle	.403 (.213)	.329 (.174)	.429 (.356)	.285 (.192)
		Final	.429 (.236)	.525 (.222)	.445 (.268)	.495 (.233)

Appendix

Phrase and no-phrase versions of lists used in Experiments 1 and 2. Chunks are in italics and critical words are underlined. Words added to form lists for Experiment 2 are marked with an asterisk. See text for explanation.

List-type	**Word 1**	**Word 2**	**Word 3**	**Word 4**	**Word 5**	**Word 6**	**Word 7**	**Word 8**	**Word 9**	**Word 10**
No-Phrase	flag	doll	*faith*	*night*	hose	film	phase*			
Phrase	flag	*good*	*faith*	*night*	*owl*	film	phase*			
No-Phrase	gear	*blue*	cheat	soap	*bird*	ink	more*			
Phrase	*baby*	*blue*	cheat	soap	*bird*	*watch*	more*			
No-Phrase	wallet	candle	she	*news*	mix	*back*	reason			
Phrase	wallet	candle	*bad*	*news*	mix	*back*	*seat*			
No-Phrase	*must*	well	*take*	relax	fin	unit	name*			
Phrase	*must*	*act*	*take*	*medicine*	fin	unit	name*			
No-Phrase	grab	*say*	pad	can	*got*	shake	thin			
Phrase	*we*	*say*	pad	can	*got*	*tired*	thin			
No-Phrase	among	*out*	coin	*ice*	tell	push	children*			
Phrase	*dry*	*out*	coin	*ice*	*cream*	push	children*			
No-Phrase	issue	*stand*	ride	send	*tall*	chew	above*			
Phrase	issue	*stand*	*proud*	send	*tall*	*tales*	above*			
No-Phrase	sing	*fly*	*cup*	rake	table	college	again*			
Phrase	*fruit*	*fly*	*cup*	*holder*	table	college	again*			
No-Phrase	smoke	*turn*	steal	ripe	*left*	know	find*	angel*		
Phrase	smoke	*turn*	*on*	ripe	*left*	*field*	find*	angel*		

No-Phrase	guess	tools	*over*	city	*down*	follow	swim	doctor*	
Phrase	guess	*fall*	*over*	city	*down*	*under*	swim	doctor*	
No-Phrase	*talk*	bone	ask	gaze	blame	*show*	knit	lot*	
Phrase	*talk*	*dirty*	ask	gaze	blame	*show*	*room*	lot*	
No-Phrase	manner	juice	*school*	pan	ruler	*jump*	packet	front*	
Phrase	manner	*high*	*school*	pan	ruler	*jump*	*rope*	front *	
No-Phrase	*right*	far	lamp	*here*	street	bake	won't*	guard*	
Phrase	*right*	*now*	lamp	*here*	*lies*	bake	won't*	guard*	
No-Phrase	band	*frame*	yet	sense	*prize*	open	nobody*	phone*	
Phrase	*door*	*frame*	yet	sense	*prize*	*cow*	nobody*	phone*	
No-Phrase	*paper*	vision	king	crowd	*boy*	wax	page	chill*	
Phrase	*paper*	*route*	king	crowd	*boy*	*scout*	page	chill*	
No-Phrase	*blood*	grow	dial	foot	van	*test*	come	cat*	
Phrase	*blood*	*bank*	dial	foot	van	*test*	*tube*	cat*	
No-Phrase	pick	*ran*	value	nail	*went*	draw	net*	urge*	unit*
Phrase	I	*ran*	value	nail	*went*	*away*	net*	urge*	unit*
No-Phrase	dear	*be*	yawn	*have*	please	long	enter	clear*	ready*
Phrase	*can't*	*be*	yawn	*have*	*fun*	long	enter	clear*	ready*
No-Phrase	seed	cord	*zone*	harp	kneel	*line*	star	hate*	board*
Phrase	seed	*time*	*zone*	harp	kneel	*line*	*drive*	hate*	board*
No-Phrase	kite	*book*	*zoo*	bean	lake	vote	study	step*	trash*
Phrase	*children's*	*book*	*zoo*	*keeper*	lake	vote	study	step*	trash*
No-Phrase	graph	*drill*	tray	*alarm*	dish	beg	color*	proud*	another*

Phrase	_fire_	_drill_	tray	_alarm_	_clock_	beg	color*	proud*	another*	
No-Phrase	number	bell	muscle	_goat_	_pass_	rotate	torn	talent*	edit*	
Phrase	number	bell	_mountain_	_goat_	_pass_	_through_	torn	talent*	edit*	
No-Phrase	palace	_ten_	oven	way	letter	_fit_	going	rude*	sand*	
Phrase	_perfect_	_ten_	oven	way	letter	_fit_	_in_	rude*	sand*	
No-Phrase	purpose	fold	cell	_hair_	bet	_spray_	wear	empty*	many*	
Phrase	purpose	fold	cell	_hair_	_piece_	_spray_	_paint_	empty*	many*	
No-Phrase	read	jar	_very_	gun	_full_	stiff	people*	knob*	role*	explain*
Phrase	read	jar	_very_	_nice_	_full_	_moon_	people*	knob*	role*	explain*
No-Phrase	_for_	alive	word	flat	_you_	house	meat*	father*	trip*	play*
Phrase	_for_	_him_	word	flat	_you_	_see_	meat*	father*	trip*	play*
No-Phrase	mineral	sew	_shot_	_day_	body	straw	large	war*	box*	oil*
Phrase	mineral	_big_	_shot_	_day_	_care_	straw	large	war*	box*	oil*
No-Phrase	craft	_face_	mind	clip	_tune_	skirt	comb*	peach*	music*	lift*
Phrase	_familiar_	_face_	mind	clip	_tune_	_pianos_	comb*	peach*	music*	lift*
No-Phrase	pile	plug	_look_	_give_	close	make	bit*	window*	trick*	sign*
Phrase	pile	_won't_	_look_	_give_	_thanks_	make	bit*	window*	trick*	sign*
No-Phrase	type	_saw_	stir	toss	church	_dug_	price	friend*	base*	not*
Phrase	_he_	_saw_	stir	toss	church	_dug_	_holes_	friend*	base*	not*
No-Phrase	_crystal_	pay	roam	_ball_	stare	they	pie	inch*	simple*	flame*
Phrase	_crystal_	_glass_	roam	_ball_	_game_	they	pie	inch*	simple*	flame*
No-Phrase	flute	_low_	stay	rat	cost	speak	hide	trend*	rough*	paste*
Phrase	flute	_low_	_voice_	rat	cost	_money_	hide	trend*	rough*	paste*

Figure Captions

Figure 1. The mean recall of extraneous words in Experiment 1 by age and list-type (right panel), together with the freed-capacity predictions of SCART (left panel) for capacity-3 (corresponding to older adults), and capacities-4 and -5.5 (corresponding to young adults). Capacity-5.5 corresponds to the mean for capacities-5 and -6.

Figure 2. The mean number of extraneous words recalled in Experiment 1 by age and list-length (right panel), together with the predictions of SCART (left panel) for capacity-3 (corresponding to older adults), and capacities-4 and -5.5 (corresponding to young adults).

Figure 3. The mean number of chunks recalled in Experiment 1 by age and list-length (right panel), together with the predictions of SCART (left panel) for capacity-3 (corresponding to older adults), and capacities-4 and -5.5 (corresponding to young adults).

Figure 4. The mean number of chunks recalled in Experiment 1 by age and list-type (right panel), together with the predictions of SCART (left panel) for capacity-3 (corresponding to older adults), and capacities-4 and -5.5 (corresponding to young adults).

Figure 5. The mean number of extraneous words recalled in Experiment 2 by age and list-type (right panel), together with the freed-capacity predictions of SCART (left panel) for capacity-3.5 (corresponding to older adults), and capacities-4 and -5.5 (corresponding to young adults).

Figure 6. The mean number of extraneous words recalled in Experiment 2 by age and list-length (right panel), together with the predictions (left panel) for capacity-3.5 (matching older adults), and capacities-4 and -5.5 (matching young adults) of SCART.

Figure 7. The mean number of chunks recalled in Experiment 2 by age and list-length (right panel), together with the predictions of SCART (left panel) for capacity-3.5 (corresponding to older adults), and capacities-4 and -5.5 (corresponding to young adults).

Figure 8. The mean number of chunks recalled in 7-word lists in Experiment 2 by age and list-type (right panel), together with the predictions (left panel) of SCART for capacity-3.5 (corresponding to older adults), and capacities-4 and -5.5 (corresponding to young adults).

Figure 9. A frequency distribution of participants with integer capacities (e.g., 1, 2, 3, 4, 5, or 6 chunks) versus off-integer capacities: Experiment 1 (left panel) and Experiment 2 (right panel).

Figure 10. Mean recall by age and list-type for critical words in length-7 lists: Experiment 1 (left panel) and Experiment 2 (right panel).

Figure 11. Mean articulation rates (left panel) and estimated phonological loop sizes (right panel) as a function of list type in Experiment 3.

Figure 1.

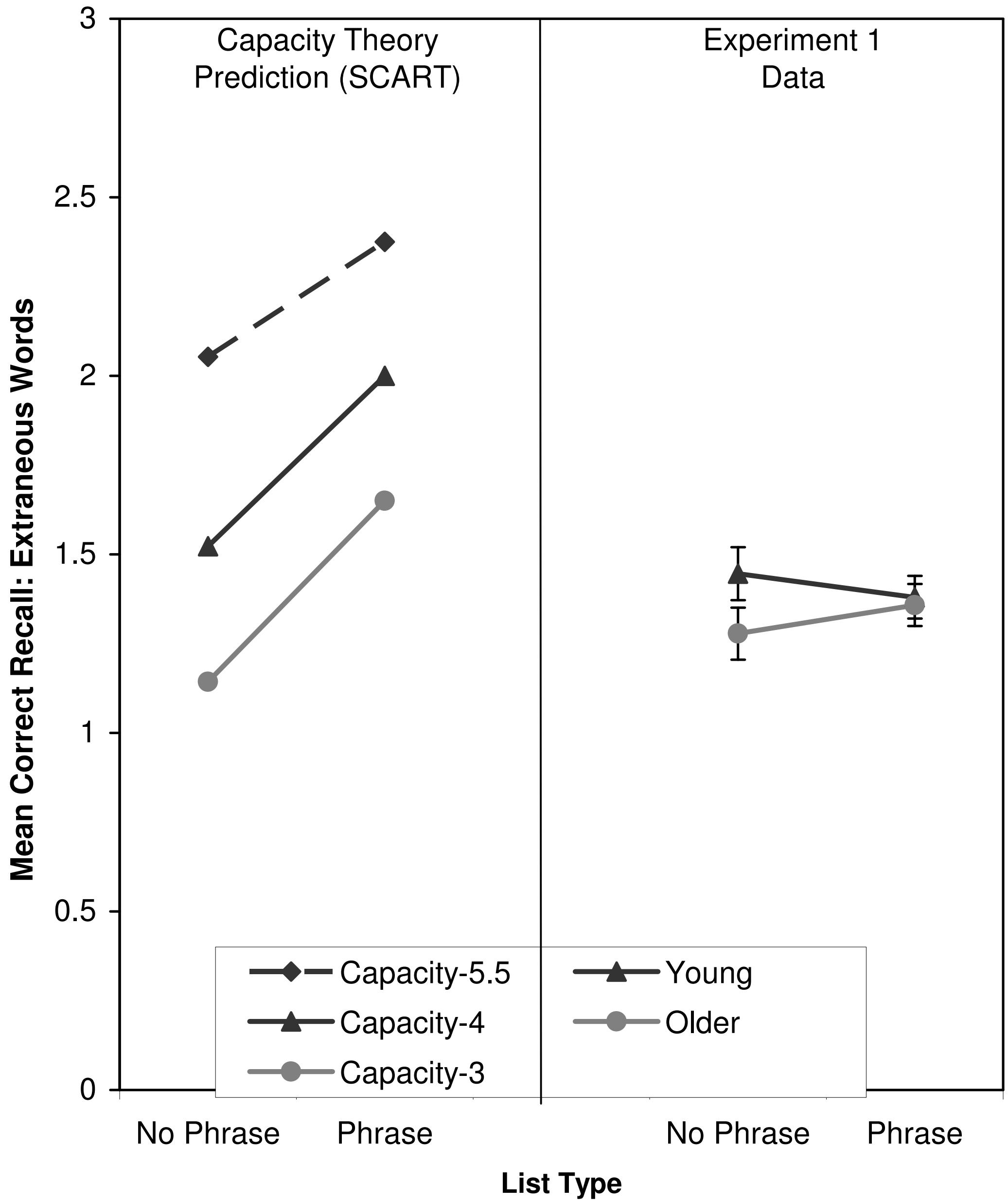

3
2.5
2
1.5
1
0.5
0
Capacity Theory Prediction (SCART)
Experiment 1 Data
Mean Correct Recall: Extraneous Words
Capacity-5.5
Capacity-4
Capacity-3
Young
Older
No Phrase
Phrase
No Phrase
Phrase
List Type

Figure 2.

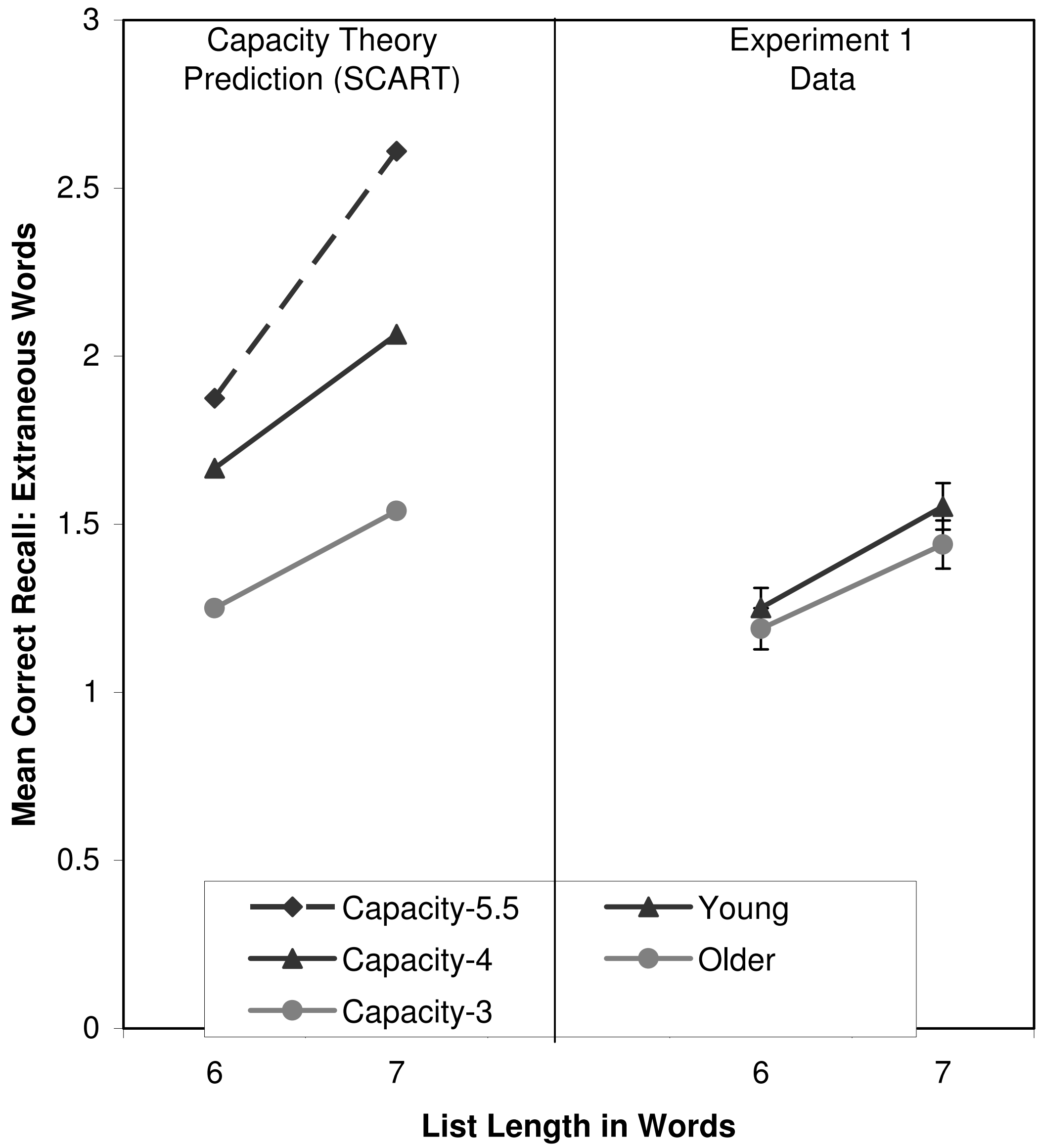
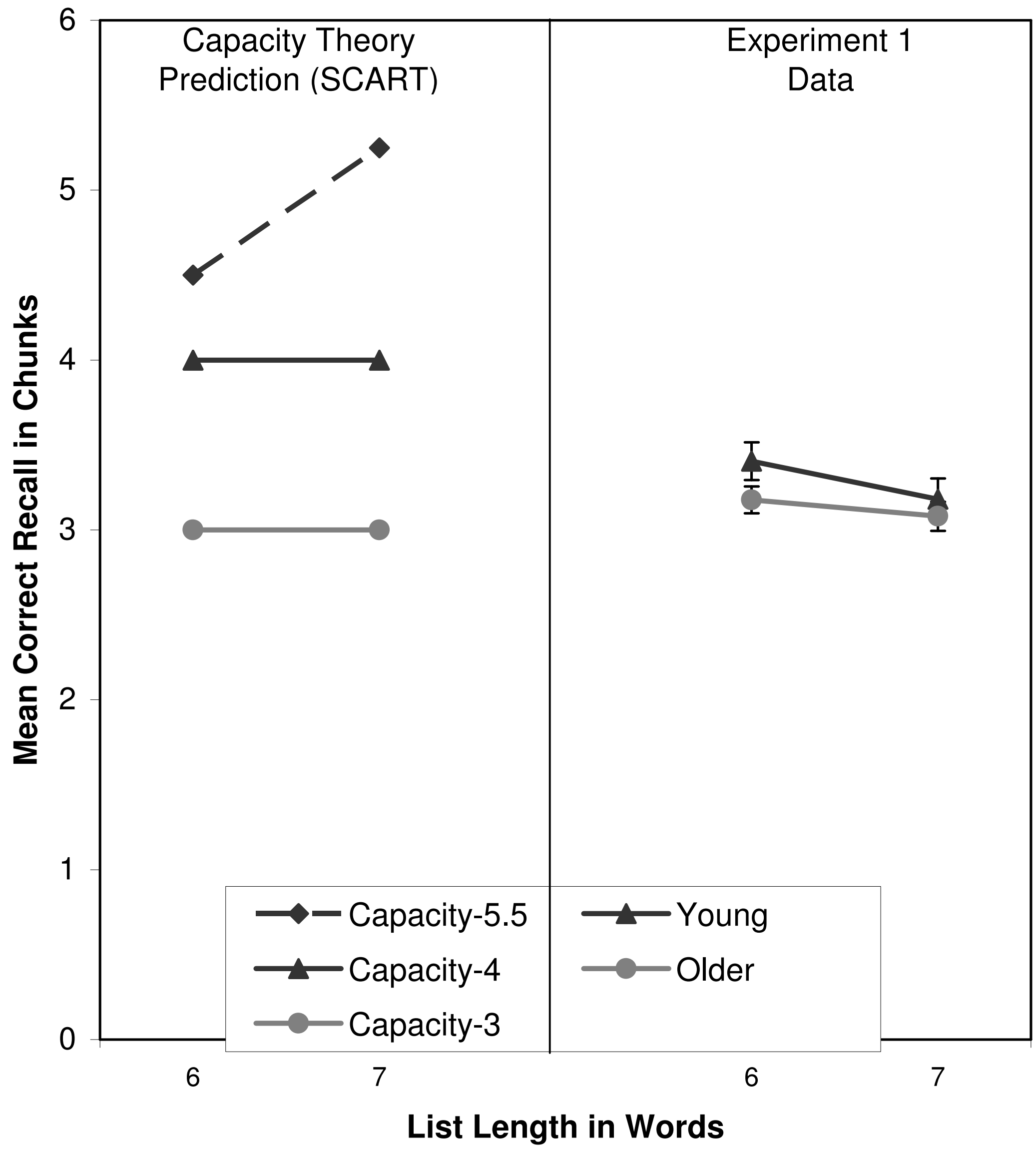
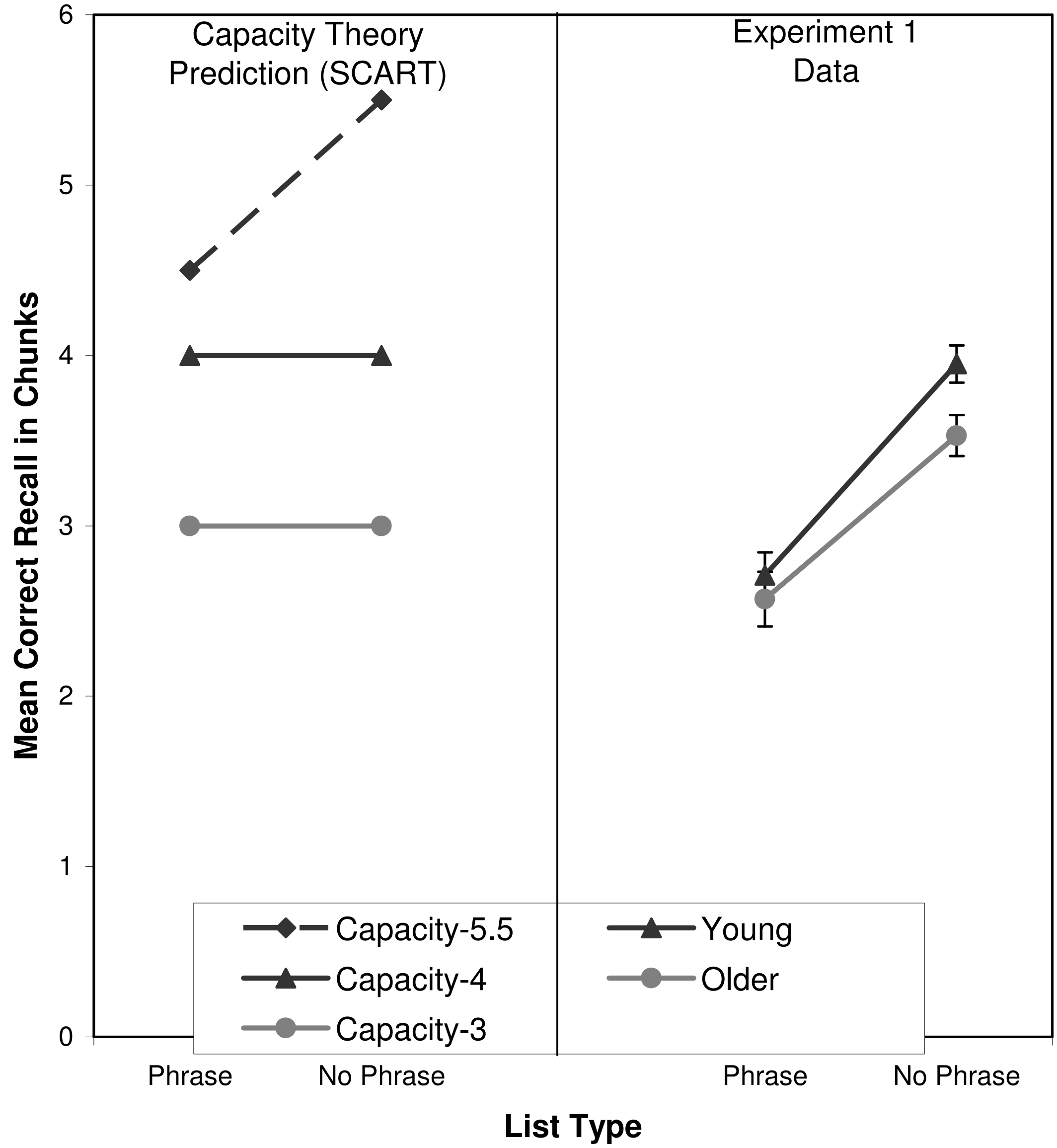

Capacity Theory
Prediction (SCART)
Experiment 1
Data
Mean Correct Recall: Extraneous Words
3
2.5
2
1.5
1
0.5
0
Capacity-5.5
Capacity-4
Capacity-3
Young
Older
6
7
6
7
List Length in Words

Figure 3.



Capacity Theory
Prediction (SCART)
Experiment 1
Data
Mean Correct Recall in Chunks
0
1
2
3
4
5
6
Capacity-5.5
Capacity-4
Capacity-3
Young
Older
6
7
6
7
List Length in Words

Figure 4.



Capacity Theory Prediction (SCART)
Experiment 1 Data
Mean Correct Recall in Chunks
0
1
2
3
4
5
6
Capacity-5.5
Capacity-4
Capacity-3
Young
Older
Phrase
No Phrase
Phrase
No Phrase
List Type

Figure 5.

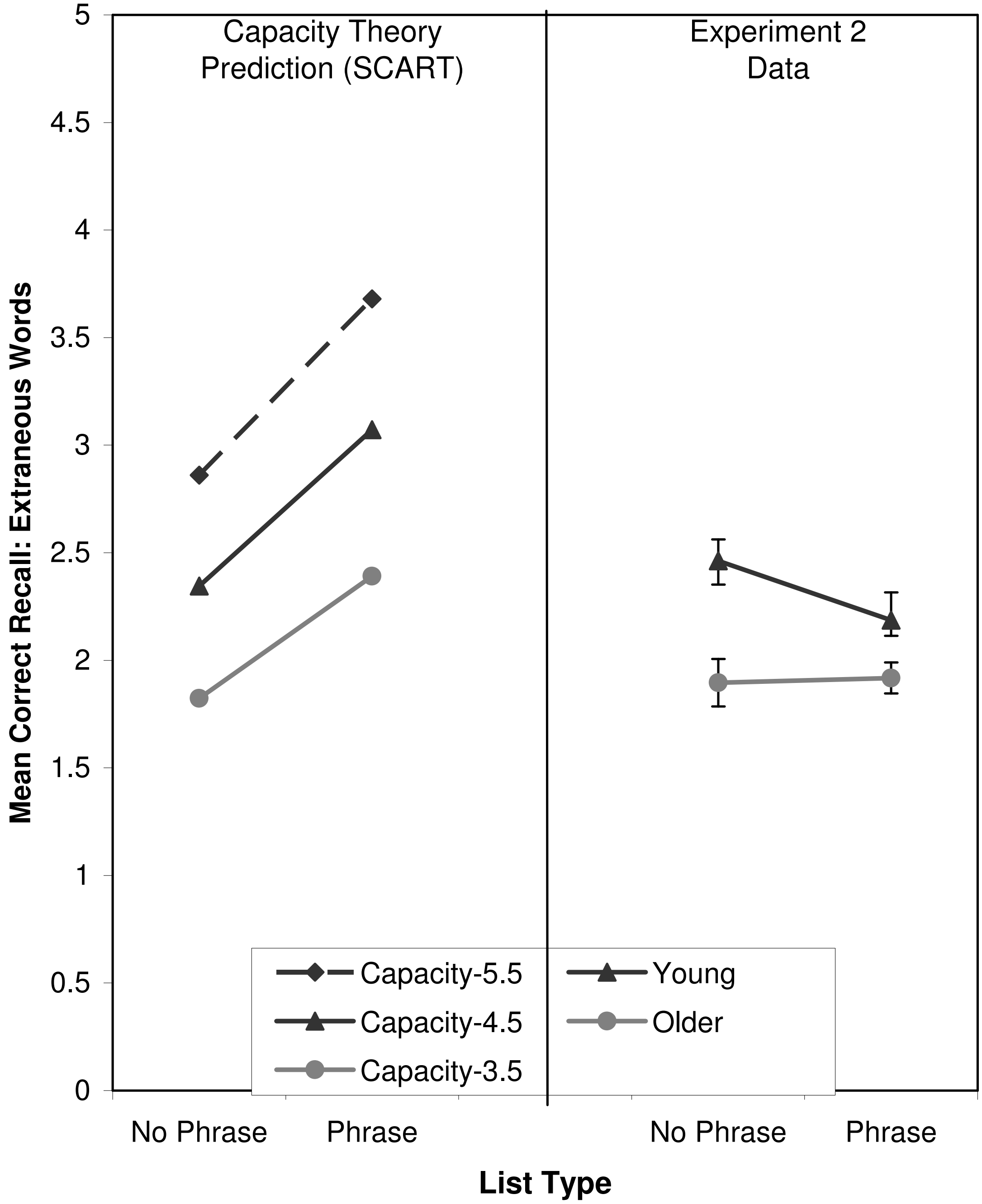

Capacity Theory
Prediction (SCART)
Experiment 2
Data
Mean Correct Recall: Extraneous Words
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
Capacity-5.5
Capacity-4.5
Capacity-3.5
Young
Older
No Phrase
Phrase
No Phrase
Phrase
List Type

Figure 6.

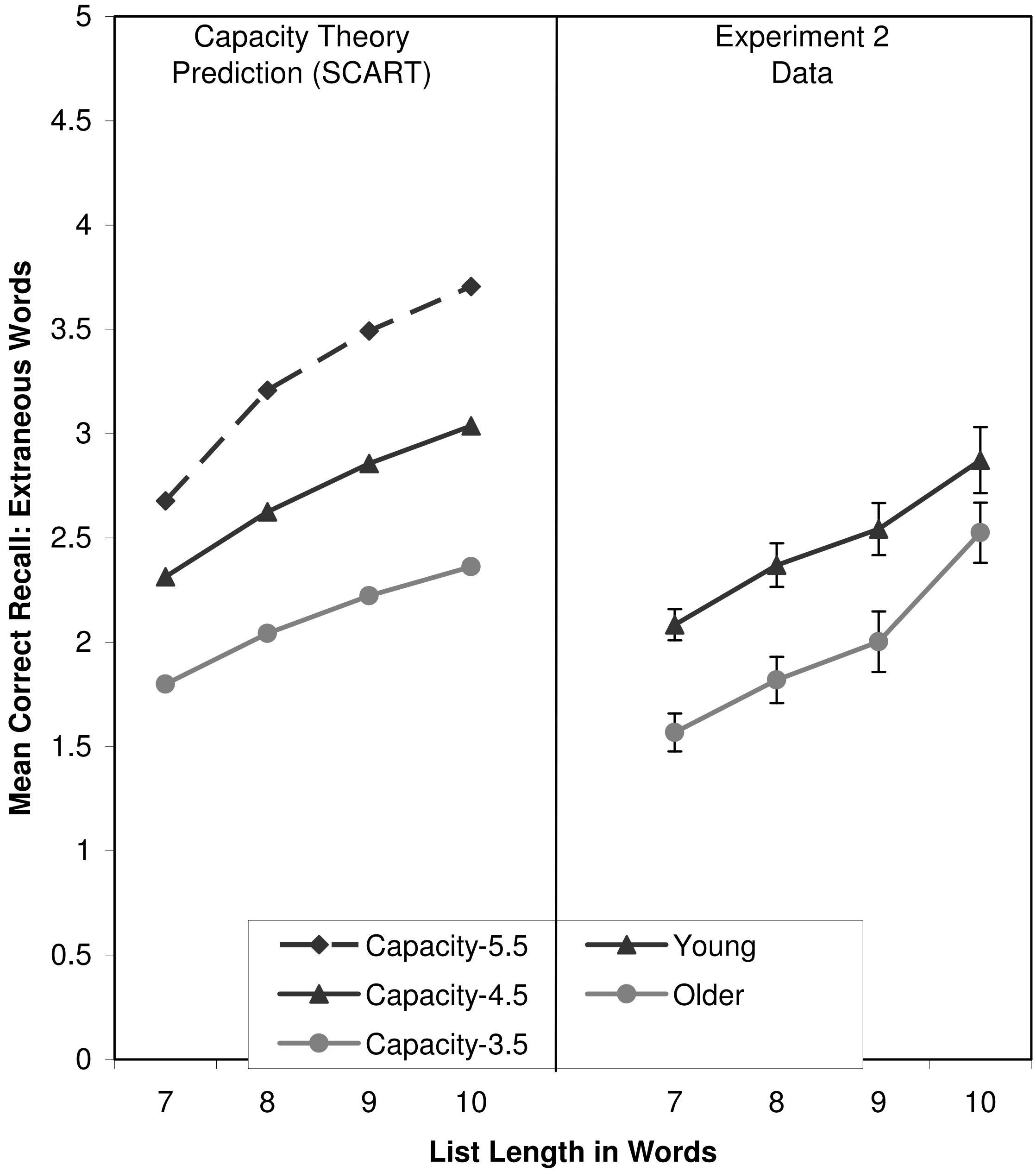

Capacity Theory
Prediction (SCART)
Experiment 2
Data
Mean Correct Recall: Extraneous Words
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
Capacity-5.5
Capacity-4.5
Capacity-3.5
Young
Older
7
8
9
10
7
8
9
10
List Length in Words

Figure 7.

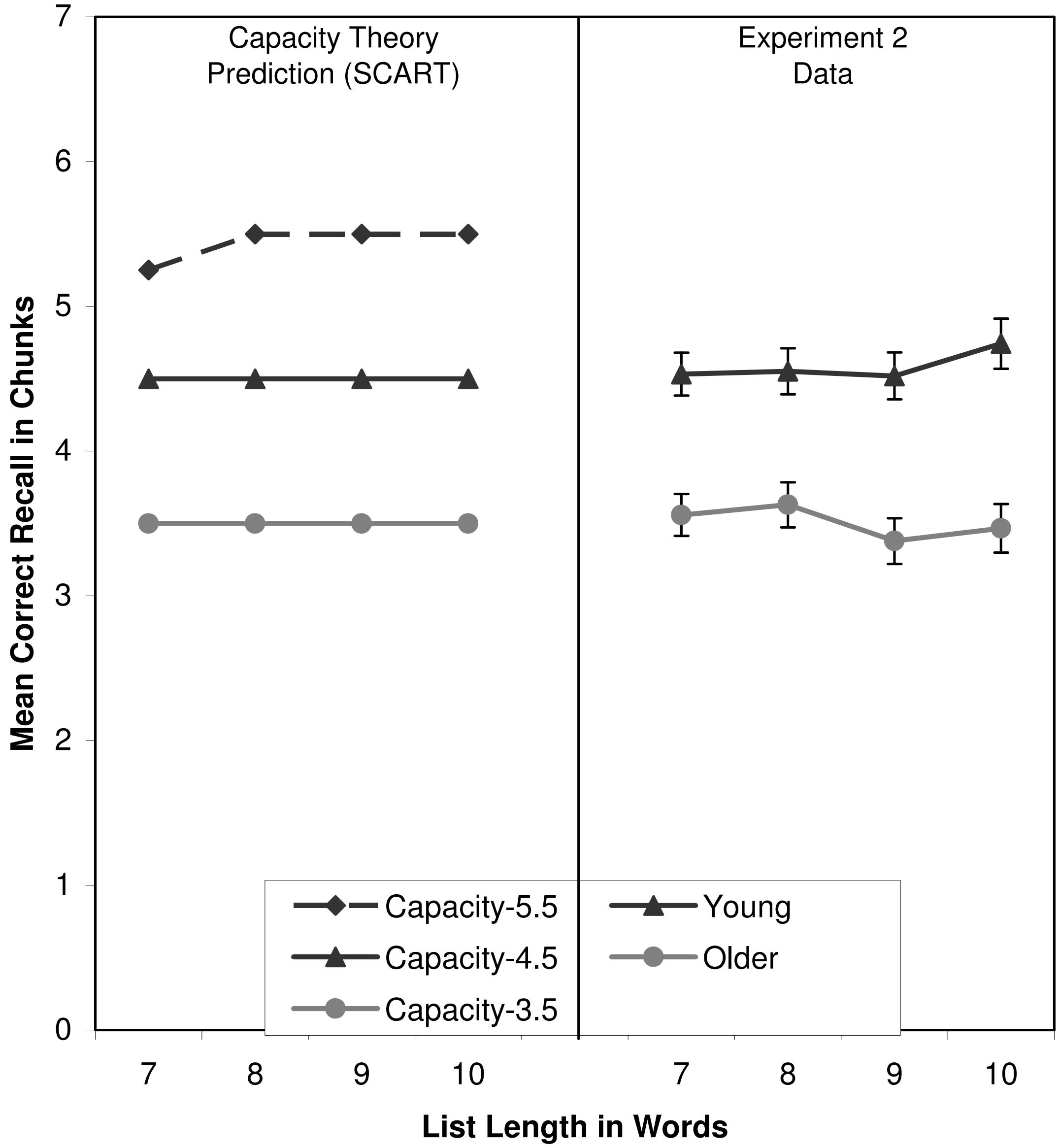
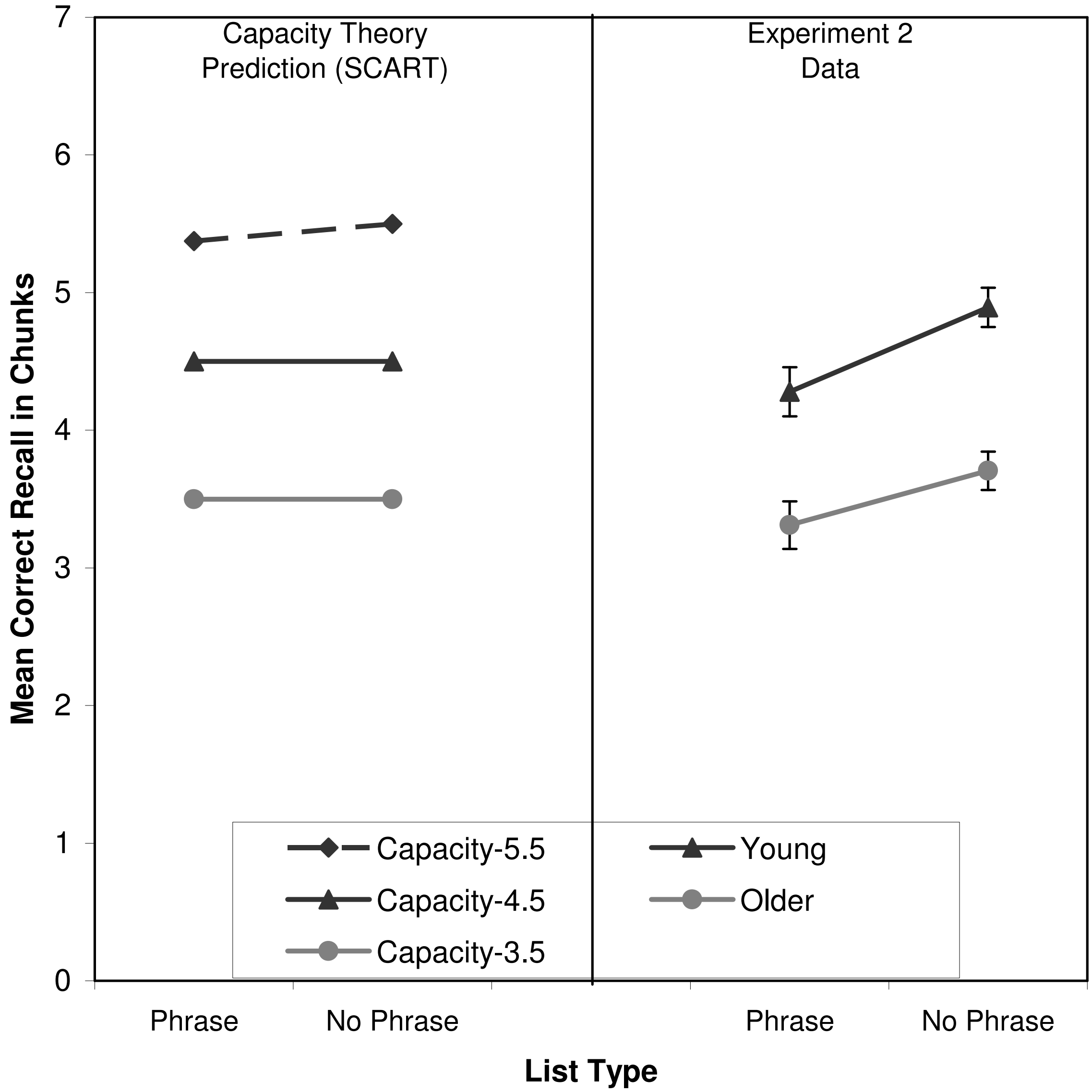

Capacity Theory Prediction (SCART)
Experiment 2 Data
Mean Correct Recall in Chunks
List Length in Words
0
1
2
3
4
5
6
7
7
8
9
10
7
8
9
10
Capacity-5.5
Capacity-4.5
Capacity-3.5
Young
Older

Figure 8.



Capacity Theory
Prediction (SCART)
Experiment 2
Data
Mean Correct Recall in Chunks
7
6
5
4
3
2
1
0
Capacity-5.5
Capacity-4.5
Capacity-3.5
Young
Older
Phrase
No Phrase
Phrase
No Phrase
List Type

Figure 9.

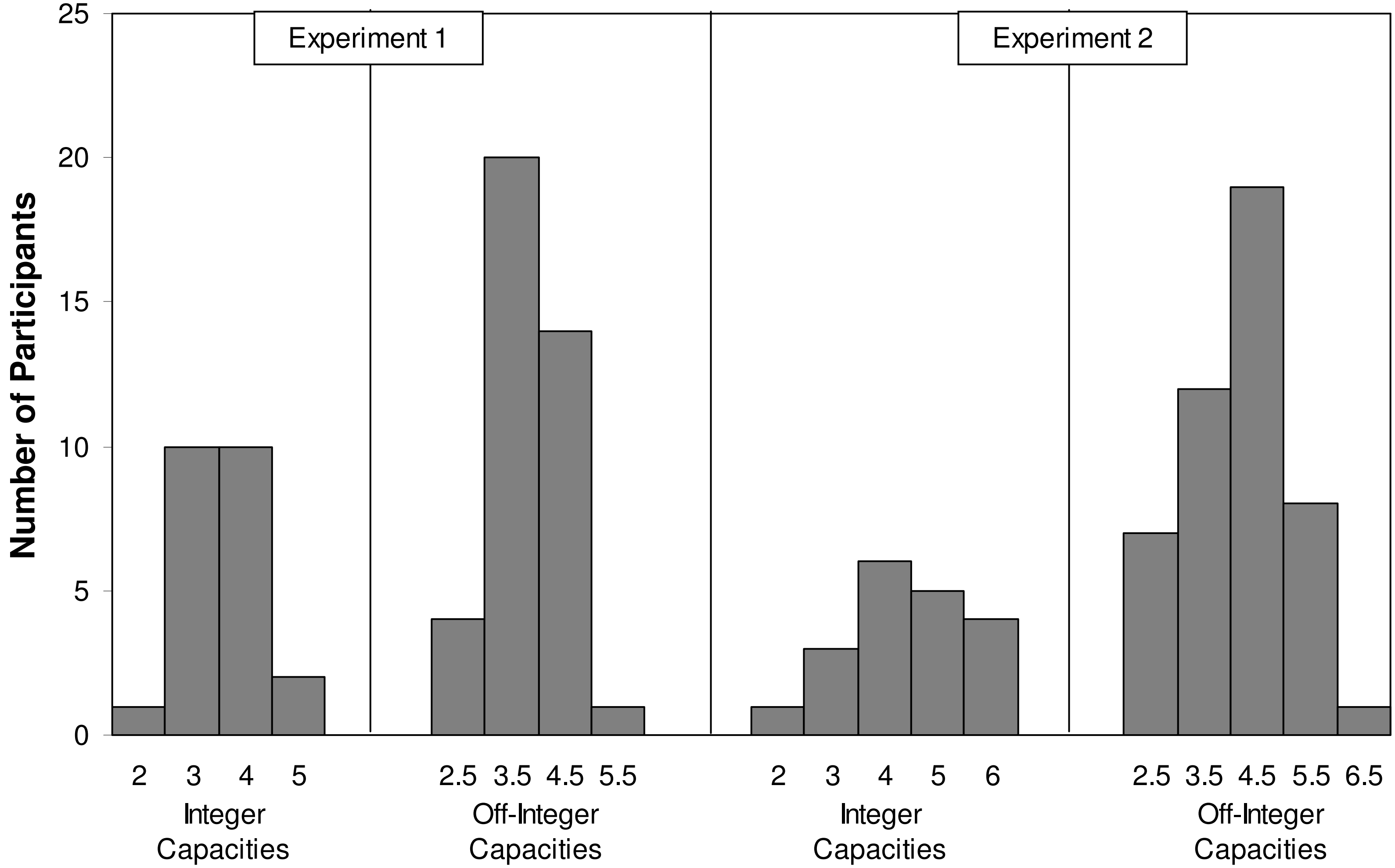

Experiment 1
Experiment 2
Number of Participants
0
5
10
15
20
25
2
3
4
5
Integer
Capacities
2.5
3.5
4.5
5.5
Off-Integer
Capacities
2
3
4
5
6
Integer
Capacities
2.5
3.5
4.5
5.5
6.5
Off-Integer
Capacities

Figure 10.

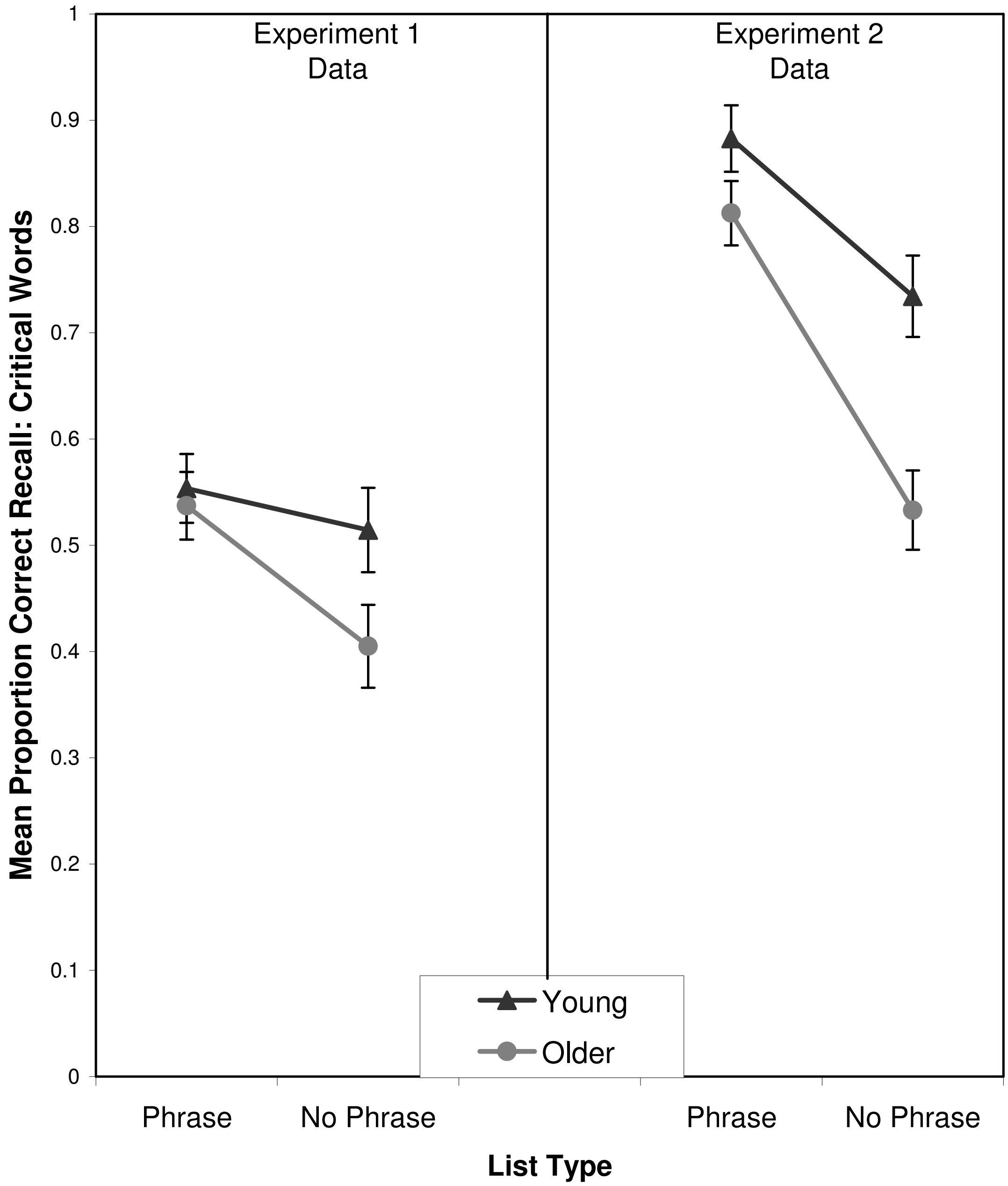
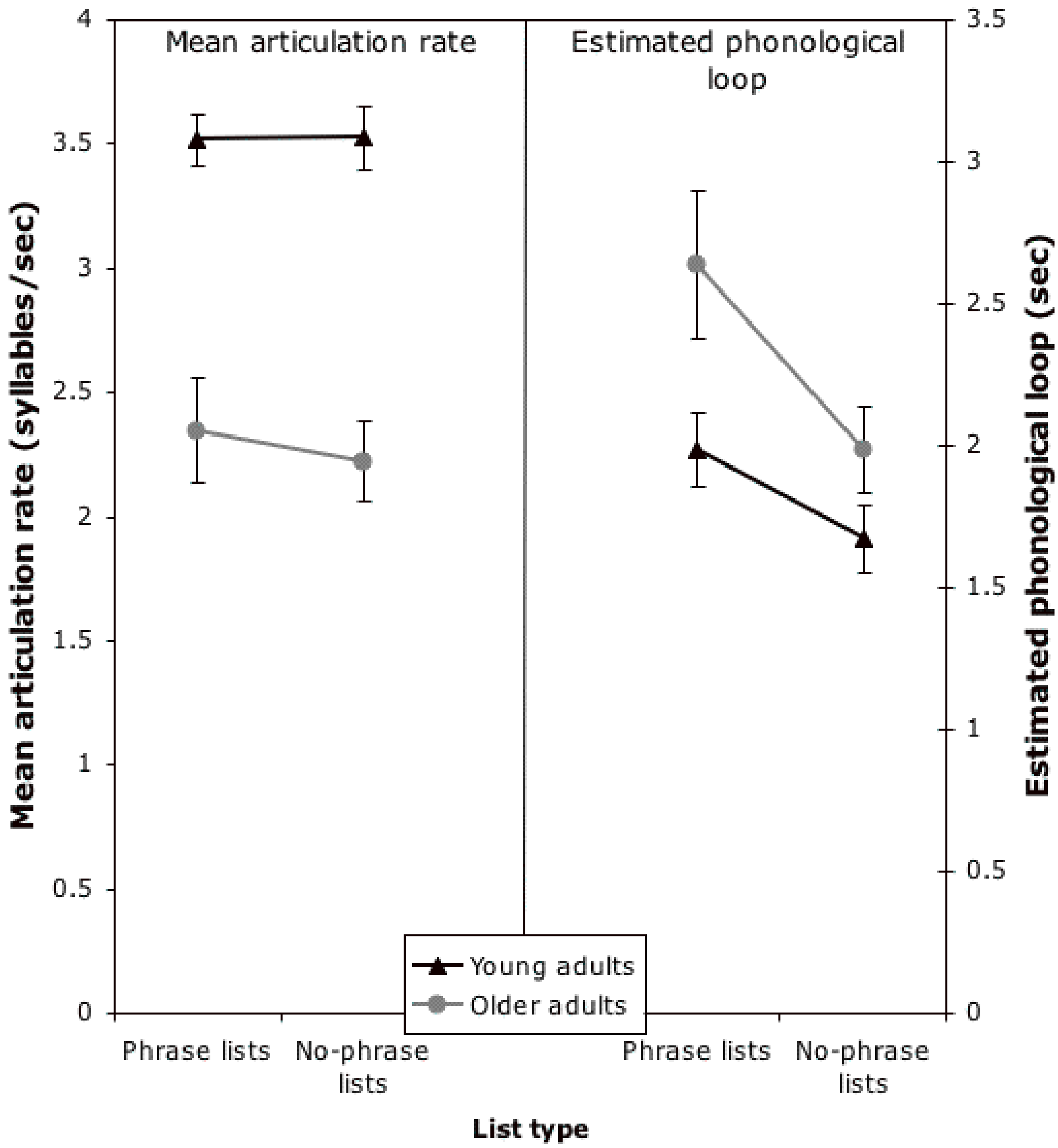

Experiment 1
Data
Experiment 2
Data
Mean Proportion Correct Recall: Critical Words
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Young
Older
Phrase
No Phrase
Phrase
No Phrase
List Type

Figure 11.


Mean articulation rate
Estimated phonological loop
Mean articulation rate (syllables/sec)
Estimated phonological loop (sec)
4
3.5
3
2.5
2
1.5
1
0.5
0
3.5
3
2.5
2
1.5
1
0.5
0
Young adults
Older adults
Phrase lists
No-phrase lists
Phrase lists
No-phrase lists
List type