THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

A list-length constraint on incidental item-to-item associations

OPEN ACCESS

A List-Length Constraint on Incidental Item-to-Item Associations

Nelson Cowan, **Kristin Donnell**, and **J. Scott Saults**
University of Missouri

Abstract

We investigated the possibility that item-to-item associations form between items concurrently included in a capacity-limited region of working memory, but not outside of that region. Many studies indicate a central capacity limit of 3 to 5 items (e.g., Cowan, 2001). Participants received lists of 3, 6, or 9 words along with an orienting task, selecting the most interesting word from each list. Consistent with expectations, a subsequent, unexpected test showed that memory of whether two words came from the same list or not was superior for 3-word lists compared to 6- and 9-word lists, which did not differ. This effect occurred even though the separation between the list positions of the two probe words was controlled across list lengths. The study demonstrates a source of implicit learning that depends upon a limited-capacity working memory faculty, a finding that should inspire further research on the function of working memory in long-term learning.

We report a new phenomenon, not intuitively obvious, that is predicted by our theory of working memory and requires explanation from whatever theory one wishes to assess. The two primary relevant theoretical assumptions are as follows: (1) several items can be present in the focus of attention at once, and (2) items that are in the focus at once tend to become associated with one another, even when no intentional effort is made to associate them (Cowan, 1999, 2001, 2005; cf. Atkinson & Shiffrin, 1968; Shiffrin & Steyvers, 1997). Based on these assumptions, we make a prediction regarding a two-phase procedure. In the first phase, lists of varying numbers of words are presented and there is an orienting task for each list in which one word is to be declared most "interesting." The second phase involves an unexpected memory test. Two words at a time, drawn from the lists, are presented again. The task is to indicate whether the words had been presented within the same list or different lists. It was expected that associations between words that had been presented within the same list would be stronger when the list was short enough so that the words would all have been likely to reside in the focus of attention at once, whereas associations would be weaker when the words were presented together in longer lists. This novel prediction should hold even with the distance between words in the list held constant. The results provide one confirmation of the theory, and are intrinsically of interest to the field even if certain alternative theories might also allow a similar prediction.

Several different procedures have been used to assess what material is included in the focus of attention (e.g., Cowan, 2011; Cowan et al., 2005; Gilchrist & Cowan, 2011; Luck &

Address Correspondence to: Nelson Cowan, Department of Psychological Sciences, University of Missouri, 18 McAlester Hall, Columbia, MO 65211. Tel. 573-882-4232; CowanN@missouri.edu.

Vogel, 1997; McElree, 1998; Oberauer, 2002; Oberauer & Hein, 2012; Reinitz & Hannigan, 2004). They differ in whether the focus of attention is said to include one, several, or a variable number of items. In line with our findings, Reinitz and Hannigan (2004) found that compound words were recombined in memory (e.g., *stargaze* and *catfish* leading to a false recognition of *starfish*) most often when the words occurred together in working memory.

Our specific prediction is based on the theory that the focus of attention typically includes 3–5 items in normal adults (Cowan, 2001); larger apparent capacities, such as the limit of about seven chunks observed by Miller (1956), occur when participants are able to engage in covert verbal rehearsal or on-line chunking (Baddeley, Thomson, & Buchanan, 1975; Broadbent, 1975; Cowan, 2001; Simon, 1974). A similar prediction might be made on the basis of a theory in which a limited number of items can reside together in a faculty outside of the focus of attention (Oberauer, 2002). In the latter theory, interference between items in a list may cause the limit (Oberauer, Lewandowsky, Farrell, Jarrold, & Greaves, 2012).

We simply documented the phenomenon of incidental learning of word-word associations by presenting lists of 3 words (within the presently assumed capacity limit), 6 words (beyond the supposed larger limit but within the limit observed by Miller, 1956), or 9 words (beyond most individuals' capacity even according to Miller). The orienting task was one in which a single list item was to be selected as most "interesting." This task ensured that words would be mentally compared, presumably by at least two words at a time being held in the focus of attention concurrently, but without any need or instruction to memorize the list. After all words had been presented once, this task was followed by an unexpected recognition memory test in which probe word pairs were to be judged to have come from the same list or different lists. We controlled the distance between the serial positions of the two probe words, the distance being equivalent across list lengths no matter whether the words came from the same list or different lists. This procedure served as a sensitive measure of the possible association between words within a list.

Method

Participants

There were 73 undergraduate students (49 female) who participated for course credit for an introductory psychology course. In order to be included in the final sample, however, the participant had to have memory data for pairs that did and did not include at least one word that had been judged most interesting in its list. This had to be the case for each of six different experimental conditions described below. The inclusion of all such conditions was left to chance and, as a result, 14 participants did not have the full complement of such data and were excluded from the analyses. This left 59 students in the final sample (41 female).

As a justification of the sample size, if participants' mean proportions correct were distributed normally with a standard deviation of .1, it would take about 50 participants to produce standard errors of about .01. It was desirable to overshoot that mark slightly in the final sample. Our obtained standard errors (see Figure 2) come close to this simple *a priori* estimate.

Design

This study comprised two consecutive phases (Figure 1). In the first phase, participants were presented with lists of 3, 6, or 9 words and, for each list, a judgment was made as to which word was most interesting. Immediately following completion of the first task, participants were given an unexpected memory task for pairs of words drawn from the lists. On each trial in this phase, two words were to be judged to come from the same list or different lists. Words were drawn from nearby serial positions (i.e., from the same word triad: both words from Serial Positions 1–3, 3–6, or 7–9), either from the same list or from two different lists of equal length. The key independent variable was the list length, and the most important dependent variable was the accuracy of the response in the memory task.

Apparatus, Stimuli, and Procedure

In a sound-attenuated room, participants were tested individually. They saw each word list with words in a single column on a computer screen. There were 36 word lists: 12 lists of each length, for a total of 216 words. The words were common, monosyllabic nouns with 2–6 letters, drawn from the MRC Psycholinguistic Database (Fearnley, 1997). They had a Kucera and Francis written frequency of 1–1207 and scored between 591 and 670 in concreteness, between 364 and 646 in familiarity, and between 459 and 667 in imagery. A few candidate words that were unusual or had multiple meanings were excluded.

Word lists with orienting task—Word lists were presented vertically (letters 11 mm tall, baselines 25 mm apart) with one word per row vertically centered on the screen. The 3-word lists were presented on the screen for 4.5 s; the 6-word lists, for 9.0 s; and the 9-word lists, for 13.5 s. Participants were to read the list aloud and then choose the word that was most interesting to them by clicking the word with the mouse before the words disappeared from the screen (Figure 1, left). The purpose of this task was to help participants focus attention on the list without intentional memorization. Each word list was presented in random order until participants had completed this task for every word list of each length.

Word-pairing memory task—Participants were given an unexpected memory task. On each trial, participants were shown two probe words from the same or different lists, one above fixation and the other below, and were to indicate by mouse click whether or not they initially appeared in the same list. Participants responded by using the mouse to select "YES" or "NO" just to the left and right of center on the probe display (Figure 1, right).

Each pair of probe words was always from the same serial position range of the list, whether or not they were from the same list. They were both from Serial Positions 1–3, both from Positions 4–6, or both from Positions 7–9. Yet, the two probe words were never from the exact same serial position in their list. This comparison task was carried out in such a manner that each word was used once, for a total of 108 trials in a randomized order.

For two-thirds of the memory trials, the probe words came from the same list. This proportion allowed perfect equivalence of serial positions tested for same-list versus different-list trials. Consider, for example, Serial Positions 1–3 in all lists. For a set of 6 lists that we might label A–F (without regard to their presentation order), the within-list pairs

included a balanced set such as A12, B13, C23, D12, E13, and F23. This left Words A3, B2, C1, D3, E2, and F1 to be used to form three between-list pairs (e.g., C1-B2, F1-A3, and E2-D3). For longer lists, the same types of balancing occurred for Serial Positions 4–6 (in 6- and 9-word lists) and for Serial Positions 7–9 (in 9-word lists).

Questionnaires—Twenty-two participants (Numbers 6 through 27) were informally questioned about their strategies after the experiment proper. The questionnaire was discontinued inadvertently when there was a change in personnel, but some useful results were nevertheless obtained.

Results

Proportion Correct

The results are reported as proportion correct judgment in the memory task. We distinguished between several types of memory trials. First, on some trials, neither probe word had been judged most interesting within its list; this occurred on 69% of the trial. On 30% of the trials, one probe word had been judged most interesting. However, on the remaining 1% of the trials, both probe words had been judged most interesting. If the participant recalled that, it was a compelling cue that the words had not come from the same pair and this cue would not depend on the formation of an association between the words; therefore, trials in which both probe words had been judged most interesting were excluded from the analysis.

Second, the probe words could come from the same list or different lists. Third, we distinguished between word pairs drawn from lists of 3, 6, and 9 words.

In an ANOVA with all of the within-subject factors mentioned above, the main effect of whether one word in the probe pair was judged most interesting did not approach significance, and neither did the main effect of whether the words came from the same pair or from different pairs; p=.99 and .23, respectively. However, as shown in Figure 2, the main effect of list length was significant, $F(2,116)=5.74$, p=.004, $\eta_p^2=.09$. Post-hoc Newman-Keuls tests indicated that memory for the pairing of probe words drawn from 3-word lists was superior to those drawn from 6- or 9-word lists, which did not differ. As the figure shows, this was a substantial effect but memory was still weak in absolute terms, even for 3-word lists. The effect is shown in Figure 3 separately for trials in which neither probe word had been judged most interesting in its list, and trials in which one probe word had been so judged. These two types of trials did not differ statistically.

The only other significant effect indicated that the response bias depended heavily on whether one of the probe words had been judged most interesting in its list. Specifically, there was a significant 2-way, crossover interaction between whether one probe word had been judged most interesting and whether the words came from the same list, $F(1,58)=45.56$, p=.00, $\eta_p^2=.44$. When neither word had been judged most interesting, performance on words from different lists (M=.60, SEM=.02) was higher than on words from the same list (M=.50, SEM=.01). In contrast, when one word had been judged most interesting, performance on words from different lists (M=.47, SEM=.02) was lower than on words from the same list

(M=.63, SEM=.02). Newman-Keuls tests indicated that, as one would expect from this crossover interaction, four of six pairwise comparisons were significant (i.e., that .63 and .60 > .50 and .47). This interaction could occur is if participants had a bias toward saying "no, not from the same list" when neither word was judged most interesting, and the opposite bias when one word was judged most interesting from its list. No other effect approached significance, p>.1 for all others.

Typically, the long-term memory for a list is most accurate for the primacy portion of the list. To remove this factor for our study, we examined memory for the first three items of each list. Examined separately for each serial position and word interest level, 58 participants had complete data. This analysis showed that the proportion correct for 3-item lists (.60) surpassed the proportions correct for 6- and 9-item lists (both .53), $F(2,114)=4.75$, $\eta_p^2=.08$; the differences between 3- and 6- and between 3- and 9-item lists both were significant by Newman-Keuls tests, with the latter two not differing. This analysis produced no effect involving word interest level.

Correlations

An additional question that might be asked is how much intra-individual consistency there is in responses. There were no significant correlations between the proportions correct for different list lengths (for each length, collapsed across pairs of words that came from the same list or different lists), or for the proportions correct between the same list length when there was or was not at least one word judged most interesting in its list.

Questionnaires

Participants who were questioned (see Methods) yielded insight into the nature of the orienting task, i.e., how they decided on the most interesting word in each list. When asked how they did so, a plurality of 41% based the judgment on the semantic qualities of the words and/or how these words related to them personally. Another 27% based it instead on esthetics: how the word would sound if pronounced (18%), how it looked or was spelled (4.5%), or both (4.5%). An additional 14% used both semantics and esthetics. Another 14% decided on the basis of which word seemed to stand out from the others, and a single remaining participant reported just using "his gut" to decide.

Discussion

Implicit associations play a very important practical role in experimental psychology; an example may be implicit associations between types of people and types of activities, which may underlie social stereotypes (e.g., Greenwald, McGhee, & Schwartz, 1998). Yet, there is relatively little understanding of how implicit associations form. The present study indicates that the formation of implicit associations is accelerated when the co-occurring items are part of a list that is only 3 items long, within what has been taken to be the adult human, core working memory capacity limit (Chen & Cowan, 2009; Cowan, 2001; Cowan, Rouder, Blume, & Saults, 2012; Oberauer, 2002), as compared to longer lists with 6 or 9 items.

The findings suggest that something about the concurrent storage of two words in a capacity-limited form of working memory promotes memory for the association between

them. In the present study, this occurred even though the orienting task was simply to indicate which word in the list was the most interesting, with no indication that a memory test was coming.

One can speculate that working memory often is used to form associations as a way to carry out a task (e.g., Cowan, 2005). The question about which word in a short list is most interesting might be answered by forming a structure of the three words, perhaps ordering them according to interest (e.g., perhaps *pig* more interesting than *wood*, in turn more interesting than *dust*); one might, for example, form a mental image of the three items (or a chunk of the whole list). Then the correct answer can be read from the structure formed. In the memory test, that structure could assist recall of the list memberships of the probe words.

Such a direct strategy in the most-interesting word task would be impossible for longer lists if a structure could not be formed quickly because of working memory constraints, and the correct answer might be obtained instead through a more piecemeal method. The first two items might be compared; the most interesting item would then be carried forward for comparison with the third item; and so on, with pairwise comparisons allowing the most interesting item to emerge. In that case, some pairs of items that were not adjacent would never be directly compared to one another and associations between them would theoretically not be formed. For example, if Item 4 in a list were ultimately judged most interesting, by this comparison method the 4–5 and 4–6 comparisons presumably would have taken place but the direct 5–6 comparison would never have taken place. In the probe word pair task for short lists, the structures formed would provide direct cues to whether the words came from the same list, but these structures would be missing for longer lists. This difference between the processing taking place for shorter versus longer lists is one way to account for the greater associative learning that took place for shorter lists.

It might be possible to explain the effects we obtained without a capacity limit but there are factors contradicting such explanations. First, suppose the successive comparison method just described was the method used for all three list lengths. This would result in a higher proportion of direct comparisons, and hence stronger associations, for words within shorter lists. That account would, however, incorrectly predict better performance for 6- than for 9-word lists; no such difference was observed. Moreover, it would predict no difference between list lengths for the first three serial positions, whereas we did find a 3-word-list advantage for those serial positions. The absence of a difference between 6- and 9-word lists also goes against any account in which the strength of word-word associations is inversely related to the list length. Finally, in a temporal distinctiveness account (e.g., Brown, Neath, & Chater, 2007) one would expect similar temporal markers for two nearby items in a list to assist in recognition that they came from the same list. Although longer lists are spread out over a longer period of time, our restriction of test trials to pairs of word that are only one or two serial positions apart should eliminate any difference based on the distinctiveness of temporal markers.

The relatively low level of performance even in the presence of new associations (as in the 3-word lists shown in Figure 2) might be explained by the difficulty in retrieving an arbitrary word-word association in the face of massive interference from other arbitrary

pairs. In ordinary learning, this interference may be overcome in several ways. First, repetition of an association multiple times may result in a memory that is stronger, not only from sheer repetition but also because information that has been presented in multiple contexts becomes less context-dependent (e.g., Barsalou, 1982; Watkins & Kerkar, 1985). Second, the arbitrariness of an association can be reduced if the participant has sufficient opportunity and motivation to think up situations in which the new association makes sense or is not arbitrary any more, and such elaborative rehearsal aids memory (Craik & Lockhart, 1972). What we observe in the present study may be the rudimentary beginning of associative memory that can form from a single co-occurrence of words in the focus of attention or, in any case, in a capacity-limited working memory, even without mnemonic intent.

Future insights could come from an examination of individual differences in working memory capacity, which have been shown to be related to retrieval of information not only in intentional conditions, but also to some extent in incidental conditions (Unsworth & Spillers, 2010). If, as Cowan et al. (2005) suggest, individuals differ in the number of items that can be encompassed in the focus of attention, and the focus-of-attention account of the present findings is correct, then participants with higher working memory capacity should tend to form incidental associations between words in longer lists than is the case for lower-capacity individuals (e.g., in lists of 4 or 5 items).

The findings thus illustrate a new paradigm that might be used to examine working memory capacity limits, the reasons for them, and their consequences for long-term memory storage.

Acknowledgments

References

Atkinson, RC.; Shiffrin, RM. Human memory: A proposed system and its control processes. In: Spence, KW.; Spence, JT., editors. The psychology of learning and motivation: Advances in research and theory. Vol. 2. New York: Academic Press; 1968. p. 89-195.

Baddeley AD, Thomson N, Buchanan M. Word length and the structure of short-term memory. Journal of Verbal Learning and Verbal Behavior. 1975; 14:575–589.

Barsalou LW. Context-independent and context-dependent information in concepts. Memory & Cognition. 1982; 10:82–93. [PubMed: 7087773]

Broadbent, DE. The magic number seven after fifteen years. In: Kennedy, A.; Wilkes, A., editors. Studies in long-term memory. Oxford, England: John Wiley & Sons; 1975. p. 3-18.

Brown GDA, Neath I, Chater N. A temporal ratio model of memory. Psychological Review. 2007; 114:539–576. [PubMed: 17638496]

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

Cowan, N. An embedded-processes model of working memory. In: Miyake, A.; Shah, P., editors. Models of Working Memory: Mechanisms of active maintenance and executive control. Cambridge, U.K: Cambridge University Press; 1999. p. 62-101.

Cowan N. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. Behavioral and Brain Sciences. 2001; 24:87–185. [PubMed: 11515286]

Cowan, N. Working memory capacity. Hove, East Sussex, UK: Psychology Press; 2005.

Cowan N. The focus of attention as observed in visual working memory tasks: Making sense of competing claims. Neuropsychologia. 2011; 49:1401–1406. [PubMed: 21277880]

Cowan N, Elliott EM, Saults JS, Morey CC, Mattox S, Hismjatullina A, Conway ARA. On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. Cognitive Psychology. 2005; 51:42–100. [PubMed: 16039935]

Cowan N, Rouder JN, Blume CL, Saults JS. Models of verbal working memory capacity: What does it take to make them work? Psychological Review. 2012; 119:480–499. [PubMed: 22486726]

Craik FIM, Lockhart RS. Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior. 1972; 11:671–684.

Fearnley S. MRC Psycholinguistic Database search program. Behavior Research Methods, Instruments, & Computers. 1997; 29:291–295.

Gilchrist AL, Cowan N. Can the focus of attention accommodate multiple separate items? Journal of Experimental Psychology: Learning, Memory, and Cognition. 2011; 37:1484–1502.

Greenwald AG, McGhee DE, Schwartz JKL. Measuring individual differences in implicit cognition: The Implicit Association Test. Journal of Personality and Social Psychology. 1998; 74:1464–1480. [PubMed: 9654756]

Luck SJ, Vogel EK. The capacity of visual working memory for features and conjunctions. Nature. 1997; 390:279–281. [PubMed: 9384378]

McElree B. Attended and non-attended states in working memory: Accessing categorized structures. Journal of Memory and Language. 1998; 38:225–252.

Miller GA. The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review. 1956; 63:81–97. [PubMed: 13310704]

Oberauer K. Access to information in working memory: exploring the focus of attention. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2002; 28:411–421.

Oberauer K, Hein L. Attention to information in working memory. Current Directions in Psychological Science. 2012; 21:164–169.

Oberauer K, Lewandowsky S, Farrell S, Jarrold C, Greaves M. Modeling working memory: An interference model of complex span. Psychonomic Bulletin & Review. 2012; 19:779–819. [PubMed: 22715024]

Reinitz MT, Hannigan SL. False memories for compound words: Role of working memory. Memory & Cognition. 2004; 32:463–473. [PubMed: 15285129]

Shiffrin RM, Steyvers M. A model for recognition memory: REM—retrieving effectively from memory. Psychonomic Bulletin & Review. 1997; 4:145–166. [PubMed: 21331823]

Simon HA. How big is a chunk? Science. 1974; 183:482–488. [PubMed: 17773029]

Unsworth N, Spillers GJ. Working memory capacity: Attention control, secondary memory, or both? A direct test of the dual-component model. Journal of Memory and Language. 2010; 62:392–406.

Watkins MJ, Kerkar SP. Recall of a twice-presented item without recall of either presentation: Generic memory for events. Journal of Memory and Language. 1985; 24:666–678.

## Word-List Trial: Click on the most interesting word.

MOUSE

BEEF

SPONGE

BAG

INN

STRAW

JAR

FRUIT

STEAK
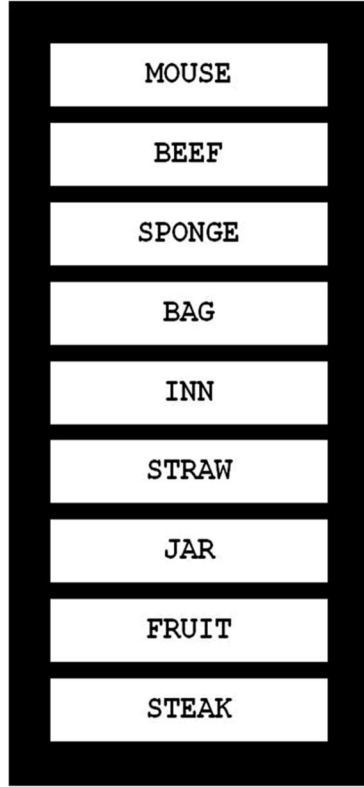
Word-Probe Trial: Were the two words from the same list?

MOUSE

NO | Same list? | YES

FLAG
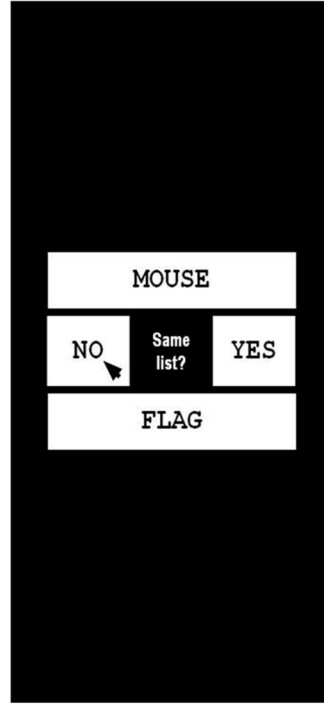
Figure 1.
An illustration of the two phases of the experiment. Left, illustration of a 9-item list from which the most interesting word was to be selected by mouse click; right, example of a probe word pair to be judged to have been presented in the same list (YES) or not (NO). Based on the list shown at the left, the correct answer must be NO because *flag* was not in the same list as *mouse* (but would have occurred in a different list, not shown). The cursor for that response is shown.
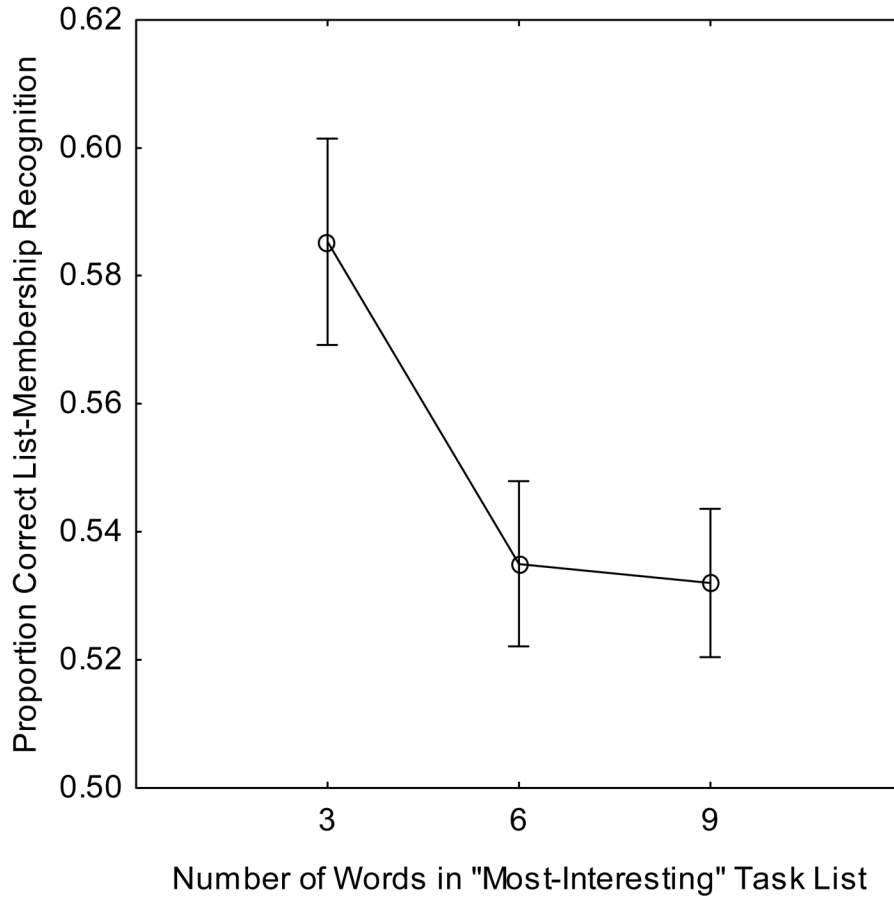
Figure 2.
Mean proportion correct judgment of probe word pairs for words taken from each list length in the "most interesting" task. In this average response, equal weight was given to trials with no word judged of most interest and with one word so judged, and equal weight to trials with probe words from the same list versus different lists. (No significant main effects of those variables emerged.) Error bars are standard errors.
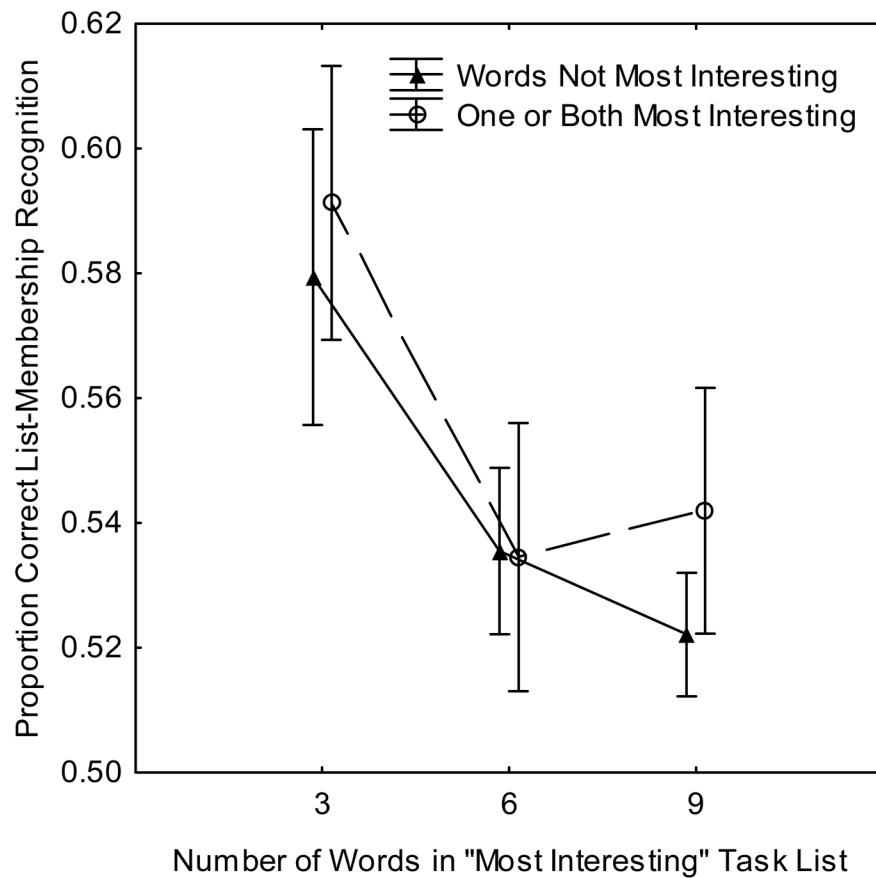
Figure 3.
Mean proportion correct judgment of probe word pairs for words taken from each list length in the "most interesting" task, shown separately for pairs for which neither word had been judged most interesting in its list, and for pairs for which one word had been judged most interesting (graph parameter). Error bars are standard errors.