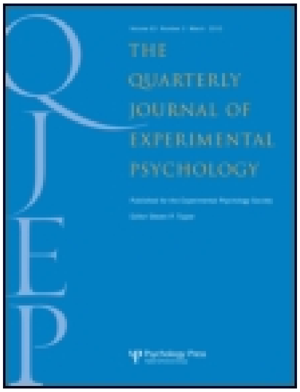

The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pqje20

Working memory capacity and retrieval limitations from long-term memory: An examination of differences in accessibility

Nash Unsworth [a] , Gregory J. Spillers [a] & Gene A. Brewer [b]
[a] Department of Psychology, University of Oregon, Eugene, OR, USA
[b] Department of Psychology, Arizona State University, Tempe, AZ, USA
Accepted author version posted online: 11 May 2012.Version of record first published: 16 Jul 2012.

To cite this article: Nash Unsworth , Gregory J. Spillers & Gene A. Brewer (2012): Working memory capacity and retrieval limitations from long-term memory: An examination of differences in accessibility, The Quarterly Journal of Experimental Psychology, 65:12, 2397-2410

To link to this article: http://dx.doi.org/10.1080/17470218.2012.690438

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2012, 65 (12), 2397–2410

Psychology Press
Taylor & Francis Group

Working memory capacity and retrieval limitations from long-term memory: An examination of differences in accessibility

Nash Unsworth[1], Gregory J. Spillers[1], and Gene A. Brewer[2]

[1]Department of Psychology, University of Oregon, Eugene, OR, USA
[2]Department of Psychology, Arizona State University, Tempe, AZ, USA

In two experiments, the locus of individual differences in working memory capacity and long-term memory recall was examined. Participants performed categorical cued and free recall tasks, and individual differences in the dynamics of recall were interpreted in terms of a hierarchical-search framework. The results from this study are in accordance with recent theorizing suggesting a strong relation between working memory capacity and retrieval from long-term memory. Furthermore, the results also indicate that individual differences in categorical recall are partially due to differences in accessibility. In terms of accessibility of target information, two important factors drive the difference between high- and low-working-memory-capacity participants. Low-working-memory-capacity participants fail to utilize appropriate retrieval strategies to access cues, and they also have difficulty resolving cue overload. Thus, when low-working-memory-capacity participants were given specific cues that activated a smaller set of potential targets, their recall performance was the same as that of high-working-memory-capacity participants.

Keywords: Working memory; Individual differences; Retrieval.

Working memory, and individual differences in working memory capacity (WMC), has long been associated with active maintenance abilities (Miyake & Shah, 1999). Recently, a number of studies have suggested that retrieval of information from long-term memory is also an important component of working memory and part of the reason for individual differences in WMC (e.g., Cowan et al., 2003; Healey & Miyake, 2009; Kane & Engle, 2000; Nelson & Goodmon, 2003; Radvanksy & Copeland, 2006; Rosen & Engle, 1997; Unsworth & Engle, 2007). In particular, recent correlational work has suggested that measures of working memory are moderately to strongly related with measures of long-term memory, and these long-term memory measures partially account for the correlation between working memory and intelligence (Mogle, Lovett, Stawski, & Sliwinski, 2008; Unsworth, 2010; Unsworth, Brewer, & Spillers, 2009). Despite initial evidence for a relation between WMC and long-term memory abilities, the reason for the

Correspondence should be addressed to Nash Unsworth, Department of Psychology, University of Oregon, Eugene, OR, 97403, USA. E-mail: nashu@uoregon.edu

http://www.psypress.com/qjep
http://dx.doi.org/10.1080/17470218.2012.690438

relation is still not fully understood. Our main goal in the present study was to better examine the reason for the relation between WMC and recall from long-term memory by examining cued and free recall with lists of categorical words.

Accessibility and the dynamics of categorical recall

An important distinction for understanding memory is that between the availability and the accessibility of items in memory (Tulving & Pearlstone, 1966). Availability refers to whether representations are effectively stored in the memory system and are available at the time of retrieval or whether the representations are unavailable due to a loss of strength due to decay or a loss of features due to overwriting or some form of interference. Accessibility refers to whether an available representation can be accessed at the time of retrieval or whether it is inaccessible due to a lack of appropriate retrieval cues. Examining categorical free and cued recall, Tulving and Pearlstone (1966) found that providing category cues to participants significantly increased the number of words recalled, and this result was primarily due to an increase in the categories recalled rather than an increase in the number of words recalled per category. Furthermore, Patterson (1972) has suggested that not only is the presence of cues at retrieval important, but the extent to which those cues specify the items of interest is also important. Thus, retrieval cues provide access to a subset of items, and the effectiveness of the retrieval cues is inversely related to the number of items subsumed under the cue (i.e., cue overload; Watkins, 1979).

This collective pattern of results can be interpreted within a search model framework in which it is assumed that two fairly independent searches are carried out to access items (Raaijmakers & Shiffrin, 1980; Rundus, 1973; Shiffrin, 1970; Unsworth, 2008). In the first phase, it is assumed that participants use an overarching general cue to sample particular retrieval cues. Next, it is assumed that after a category cue is sampled, items subsumed under the category cue are sampled. This hierarchical search model predicts that participants should generate clusters or bursts of items that are highly similar in free recall (e.g., Mandler, 1975, 2011). In cued recall, where the category labels are presented at recall, there is no need to search for the categories, and thus recall should proceed much like that within category searches. Furthermore, the increase in the number of items recalled in cued recall over free recall occurs because participants no longer have to search for the category labels and, thus, only within category searches limit the total number of items recalled (Tulving & Pearlstone, 1966).

Additionally, hierarchical search models of this type also account for inter-response times (IRTs) both within and between categories (Wixted & Rohrer, 1994). For instance, Pollio, Richards, and Lucas (1969) found that IRTs within a cluster were relatively fast, whereas IRTs between clusters were much slower (see also Patterson, Meltzer, & Mandler, 1971). Although much early work has examined these issues, less work has examined these issues in terms of individual differences in working memory. Indeed, Mandler (2011) recently noted that “further research is necessary to investigate organization theory’s relevance to phenomena such as individual differences, working memory, and attentional capacities” (p. 234).

Working memory capacity and recall from long-term memory

As noted previously, recent work has suggested that individual differences in WMC as measured by complex span tasks not only reflect differences in active maintenance abilities (Engle & Kane, 2004) but also reflect differences in the ability to retrieve information from long-term memory (Unsworth & Engle, 2007). To account for these differences, we (Unsworth, 2007; Unsworth & Engle, 2007) suggested a search model similar to that of Shiffrin (1970) in which there are both directed and random components to the overall search process (Shiffrin, 1970; Shiffrin & Atkinson, 1969). Directed control processes include setting up a retrieval plan, selecting and generating appropriate cues to search memory with, and various monitoring strategies and decisions to continue searching or not. The random component refers to the probabilistic nature of the search process in which a subset of information is activated by the cues (the search set), and representations are subsequently sampled and recovered from this subset (Raaijmakers & Shiffrin, 1980; Shiffrin, 1970).

In terms of the relation between WMC and long-term memory recall, we have suggested that individual differences in WMC largely reflect differences in the directed components of the search process (Rosen & Engle, 1997; Unsworth & Engle, 2007). In particular, we have suggested that low-WMC individuals are poorer at generating internal retrieval cues to focus their search sets than are high-WMC individuals. Accordingly, this framework predicts that providing low-WMC individuals with the appropriate cues (and obviating the need to internally generate cues) should equate high- and low-WMC individuals in recall. Some recent work has provided evidence consistent with this hypothesis in that individual differences in WMC were drastically reduced in cued recall compared to free recall (Unsworth, 2009b), but there were no changes when comparing specific and nonspecific cues in cued recall (Unsworth, 2009a). Thus, it remains unclear whether differences in WMC and long-term memory recall are partially due to differences in accessibility in terms of differences in the ability to generate internal retrieval cues.

Our goal in the present study was to better test whether individual differences in WMC and recall from long-term memory are partially due to differences in accessibility. This view predicts that providing low-WMC individuals with appropriately focused retrieval cues should mitigate their retrieval deficits, thus bringing them up to the level of high-WMC individuals. Note, we are not suggesting that encoding differences do not matter. On the contrary, we and others have suggested that part of the reason for the relation between WMC and recall from long-term memory is due to strategic encoding factors (e.g., Bailey, Dunlosky, & Kane, 2008; Cokely, Kelley, & Gilchrist, 2006; Unsworth & Spillers, 2010). However, in the present study we primarily focus on potential differences in accessibility of information in long-term memory.

EXPERIMENT 1

The purpose of Experiment 1 was to examine WMC differences in long-term memory recall in terms of differences in accessibility. Participants were presented with lists of categorical words. The words were presented either randomly or blocked by category. At recall, participants recalled either freely or in the presence of one of the category labels as a cue (Incisa della Rocchetta & Milner, 1993). If WMC differences in recall are partially due to differences in accessibility of retrieval cues, then providing participants with retrieval cues should alleviate their recall deficits, and the free versus cued recall manipulation should interact with WMC such that high- and low-WMC individuals should have similar performance under cued but not free recall conditions. Specifically, under free recall conditions, high-WMC individuals should recall more items, recall more categories (clusters), and have faster IRTs than low-WMC individuals. However, under cued recall conditions, high- and low-WMC individuals should perform equivalently. Furthermore, if WMC differences in recall are due, in part, to problems with organizational abilities at encoding, then providing participants with the material blocked by category should boost performance and interact with WMC such that high- and low-WMC individuals demonstrate similar performance under blocked encoding but not random encoding conditions.

Method

Participant and WMC screening

Participants were recruited from the participant pool at the University of Georgia. Individuals were selected based on a z score composite of three complex span tasks. Only participants falling in the upper (high-WMC individuals) and lower (low-WMC individuals) quartiles of the composite distribution were selected. Participants were 30 high-WMC individuals (z-WMC = 0.84, SD = 0.18) and 30 low-WMC individuals (z-WMC = –1.01, SD = 0.52), as determined by the composite measure. Note that we utilized an extreme groups design (rather than a continuous design) because we were interested in examining whether a relationship exists between WMC and aspect of recall from long-term memory (rather than estimate the magnitude of the relation) in the most cost-efficient manner. Given that the distributions for these measures are normally distributed, and prior work has found similar results using both extreme groups and the full range of participants (Unsworth, 2007, 2009c), one would expect the current results to generalize when examining the full range of participants and a more diverse set of tasks.

Operation span. Participants solved a series of math operations while trying to remember a set of unrelated letters. At recall, letters from the current set were recalled in the correct order by clicking on the appropriate letters (see Unsworth, Heitz, Schrock, & Engle, 2005). For all of the WMC measures, items were scored if the item was correct and in the correct position. There were 75 trials. The score was the number of correct items in the correct position.

Reading span. Participants were required to read sentences while trying to remember a set of unrelated letters. At recall, letters from the current set were recalled in the correct order by clicking on the appropriate letters. There were 75 trials. The score was the number of correct items in the correct position (see Unsworth, Redick, et al., 2009, for more details).

Symmetry span. Participants were required to recall sequences of red squares within a matrix while performing a symmetry-judgement task. At recall, participants recalled the sequence of red-square locations in the preceding displays, in the order they appeared, by clicking on the cells of an empty matrix. There were 42 trials. The score was the number of correct items in the correct position (see Unsworth, Redick, et al., 2009, for more details).

Composite score

In both experiments, participants were prescreened for WMC based on a composite of the three WMC tasks. Only participants falling in the upper and lower quartiles from this composite distribution completed the recall tasks. At the time of this study, the three complex span tasks were strongly correlated with one another (operation span–reading span r = .59, operation span–symmetry span r = .50, reading span–symmetry span r = .45) in our overall distribution with roughly 1,100 participants. These results are consistent with much prior research (Redick et al., in press). For the composite score, scores for each of the three complex span tasks were z-transformed for each participant. These z scores were then averaged together, and quartiles were computed from the averaged distribution.

Categorized recall procedure

In this task, participants were given four lists of 24 words each. Each list contained four different categories with 6 words per category. Words and category labels were from Murdock (1976) excluding the five most frequent words for each category. Categories included: body parts, parts of a house, types of boats, spices, kitchen utensils, vegetables, birds, occupations, fruits, insects, fabrics, tools, animals, colours, musical instruments, and trees. At presentation, each word was presented alone in the centre of the screen for 2 s each. Following a 1-s interval, the next word was presented. In the random encoding condition, the 24 words were presented randomly without any category labels. In the blocked encoding condition, the words were presented such that first the category label was presented alone for 3 s followed by each of the six words (presented alone for 2 s each) from that category (Incisa della Rocchetta & Milner, 1993). In the free recall condition, participants had 2 min to recall as many of the 24 words from the most recently presented list as possible in any order they wished by typing the words and pressing ENTER to record their response. In the cued recall condition, participants were presented with one of the category labels and were instructed to recall those items from the list that matched the category label. Participants had 30 s to recall as many items matching the category cue as possible in any order they wished by typing the words and pressing ENTER to record their response. Pilot work suggested that 30 s was enough time for participants to reach asymptotic levels of recall. After 30 s, the next category cue was presented, and participants were again instructed to recall the words matching the new category label. Thus, within a 2-min recall period, participants had 30 s to recall from each of four different categories. All participants performed each list. All participants first received the random–free list, followed by the random–cue list, the block–free list, and finally the block–cue list. For each list, participants first learned that list and then recalled the list before moving onto the next list. Note that the random lists were presented before the blocked lists in the hopes of lessening the possibility of participants guessing that the lists were composed of categorically related items. Prior work (e.g., Cohen, 1966) has suggested that if participants are aware of the categorical nature of the lists, there are only slight differences between blocked and random conditions. Thus, to maximize potential differences, the random lists were always presented before the blocked lists.

Results

Proportion of words recalled

First we examined the proportion of words recalled with a 2 (encoding: blocked vs. random) × 2 (retrieval: cued vs. free) × 2 (WMC: high vs. low) analysis of variance (ANOVA) with encoding and retrieval conditions as within-subject factors and WMC as a between-subjects factor. Shown in Table 1 are the proportions of words recalled for each of the four lists as a function of WMC. As can be seen, high-WMC individuals recalled more words on average than low-WMC individuals ($M = .67$, $SE = .02$ vs. $M = .57$, $SE = .02$), $F(1,\ 58) = 12.72$, $MSE = .04$, $p < .01$, $\eta_p^2 = .18$. A higher proportion of words were recalled under cued recall than under free recall ($M = .65$, $SE = .01$ vs. $M = .58$, $SE = .01$), $F(1,\ 58) = 34.57$, $MSE = .01$, $p < .01$, $\eta_p^2 = .37$. As shown in

Table 1. *Proportion correct as a function of conditions and WMC for Experiment 1*

	Condition			
WMC	*Random–free*	*Random–cue*	*Block–free*	*Block–cue*
High	.71 (.02)	.64 (.02)	.59 (.03)	.73 (.02)
Low	.57 (.03)	.59 (.03)	.47 (.03)	.64 (.02)

Note: Values in parentheses represent one standard error of the mean. WMC = working memory capacity.

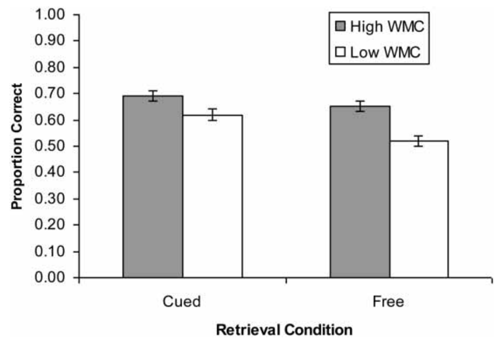

Figure 1. *Proportion correct as a function of working memory capacity and retrieval condition. Error bars represent one standard error of the mean. WMC = working memory capacity.*

Figure 1, there was a WMC × Retrieval Condition interaction, $F(1,\ 58) = 6.35$, $MSE = .01$, $p < .05$, $\eta_p^2 = .10$, such that cued recall boosted performance for low-WMC individuals to a greater extent than it boosted performance for high-WMC individuals. Specifically, under cued recall conditions, low-WMC individuals increased their proportion recalled by .10 ($SE = .02$) whereas high-WMC individuals only increased their proportion recalled by .04 ($SE = .01$), $t(58) = 2.52$, $p < .05$, $\eta^2 = .10$. Thus, high- and low-WMC individuals demonstrated a greater difference under free recall than under cued recall. Although cued recall boosted recall performance for low-WMC individuals, it did not bring them up to the level of high-WMC individuals. That is, even under cued recall conditions, high-WMC individuals still outperformed low-WMC individuals, $t(58) = 2.70$, $p < .01$, $\eta^2 = .11$. The only other effect to reach conventional levels of significance was the Encoding × Retrieval Condition interaction, $F(1, 58) = 39.60$, $MSE = .01$, $p < .01$, $\eta_p^2 = .41$. This interaction suggested that the difference between cued and free recall conditions was strongest when items were blocked at encoding rather than when items were mixed randomly at encoding. This interaction is consistent with much prior work on encoding specificity (e.g., Dong, 1972; Tulving, 1983; Unsworth, 2009b) suggesting that the difference between cued and free recall is greatest when encoding and retrieval conditions match. Although the three-way interaction with WMC was not significant, it is interesting that high-WMC individuals recalled fewer items in the random–cue condition than in the random–free condition, $t(29) = 3.32$, $p < .01$, but low-WMC individuals showed no difference between these conditions ($t < 1$).

Number of categories recalled for free recall

Next, we examined the number of categories (clusters) recalled under free recall conditions only given that under cued conditions all items will be from the same category. Presumably this reflects the accessibility of higher order units within a cue-dependent forgetting framework (Tulving & Pearlstone, 1966). In these analyses, a new category (cluster) was deemed to be recalled when at least one word was recalled from the category prior to any other words being recalled. A category was deemed as repeated if recall from one category intervened between the recall of words from the same category. Thus, the new category measure should give an indication of overall accessibility of categories, whereas the repeated category measure should provide an indication of the number of times participants resampled the same category. Shown in Table 2 are the number of new and repeated categories as function of encoding condition and WMC.

First we examined new category recall with a 2 (encoding: blocked vs. random) × 2 (WMC: high vs. low) ANOVA. The results suggested a main effect of WMC, $F(1, 58) = 13.48$, $MSE = 0.85$, $p < .01$, $\eta_p^2 = .19$, such that high-WMC individuals recalled more categories than low-WMC individuals ($M = 3.65$, $SE = 0.12$ vs. $M = 3.03$, $SE = 0.12$). That is, high-WMC individuals tended to have better access to categories than low-WMC individuals. The only other effect to approach conventional levels of significance was an effect of encoding condition, $F(1, 58) = 3.37$, $MSE = 0.71$, $p < .08$, $\eta_p^2 = .06$, suggesting that participants recalled slightly more categories in the random condition than in the blocked condition ($M = 3.48$, $SE = 0.11$ vs. $M = 3.2$, $SE = 0.12$). However, this effect seemed to be driven by a single participant who did not generate a single cluster of categorically related items in the blocked condition. Eliminating this participant from the analysis suggested that the effect of encoding condition was no longer significant ($p > .13$), but the main effect of WMC remained ($p < .01$).

A similar 2 (encoding: blocked vs. random) × 2 (WMC: high vs. low) ANOVA for repeated categories suggested only an effect of encoding condition, $F(1, 58) = 12.77$, $MSE = 0.35$, $p < .01$, $\eta_p^2 = .18$, suggesting that participants were more likely to repeat categories in the random condition than in the block condition ($M = 0.52$, $SE = 0.10$ vs. $M = 0.13$, $SE = 0.04$). Thus, high-WMC individuals accessed categories better than low-WMC individuals, but there was no difference in the tendency to resample categories.

Table 2. *Number of new and repeated categories recalled as a function of encoding condition and WMC for Experiment 1*

	Measure			
WMC	*Random–new*	*Random–repeated*	*Block–new*	*Block–repeated*
High	3.87 (0.16)	0.50 (0.14)	3.43 (0.16)	0.07 (0.06)
Low	3.10 (0.16)	0.53 (0.14)	2.97 (0.16)	0.20 (0.06)

Note: Values in parentheses represent one standard error of the mean. WMC = working memory capacity.

Number of words per category for free recall

In addition to examining the number of categories recalled, we also examined the number of words recalled per category in succession in free recall only. That is, we examined how many words on average were recalled in succession (within a cluster) from a given category. This should provide an indication of accessibility of words within a category. A 2 (encoding: blocked vs. random) × 2 (WMC: high vs. low) ANOVA examined the number of words per category for free recall. The only effect to reach significance was an effect of WMC, $F(1, 58) = 14.80$, $MSE = 0.60$, $p < .01$, $\eta_p^2 = .20$, suggesting that high-WMC individuals recalled more words per category than low-WMC individuals ($M = 3.60$, $SE = 0.10$ vs. $M = 3.06$, $SE = 0.10$).

Inter-response times

In our final set of analyses, we examined inter-response times (IRTs) for items recalled in free recall both between categories and within categories as a function of encoding condition and WMC. IRTs were measured as the difference between the first key stroke on item n and the first key stroke on item $n+1$. Note, only IRTs associated with new clusters are reported. Including IRTs associated with repeated clusters led to identical results. Shown in Table 3 are the corresponding IRTs. These results were examined with a 2 (encoding: blocked vs. random) × 2 (IRT: within vs. between) × 2 (WMC: high vs. low) ANOVA. The results suggested a main effect of within- versus between-category IRTs, $F(1, 58) = 55.32$, $MSE = 83{,}080{,}000$, $p < .01$, $\eta_p^2 = .49$, such that within-category IRTs were much faster than between-category IRTs ($M = 3.31$ s, $SE = 0.20$ vs. $M = 7.03$ s, $SE = 0.45$). Additionally, there was a main effect of WMC, $F(1, 58) = 9.37$, $MSE = 12{,}890{,}000$, $p < .01$, $\eta_p^2 = .14$, suggesting that high-WMC individuals had faster IRTs than low-WMC individuals ($M = 4.43$ s, $SE = 0.34$ vs. $M = 5.90$ s, $SE = 0.34$). None of the other effects reached conventional levels of significance (all ps > .20). Interestingly, within-category IRTs in free recall were not significantly different from IRTs in category cued recall, $t(59) = 1.46$, $p > .15$, suggesting that how participants accessed items once in a category in free recall was the same as how they accessed items when given the appropriate cue in cued recall.[1]

Discussion

The results from Experiment 1 were consistent with hierarchical search models of categorized free recall. The results suggested that in free recall, participants tended to recall items in categorical clusters before switching to new categorical clusters of words and that IRTs within a category were much faster than IRTs between categories. Furthermore, once category labels were provided as cues, performance increased, and IRTs within categories were not different from IRTs for category cued recall.

In terms of WMC, the results are consistent with the notion that high- and low-WMC individuals partially differ in accessibility. High-WMC individuals recalled more words, more categories, and more words per category than low-WMC individuals in free recall. High-WMC individuals also had faster IRTs both within and between categories in free recall. Thus, high-WMC individuals were better at accessing the category cues and better at accessing items within the categories than low-WMC individuals in free recall. Once category labels were provided as cues in cued recall, the difference between high- and low-WMC individuals was reduced due to a boost in overall performance for low-WMC individuals. This suggests that part of the difference in recall between high- and low-WMC individuals was in the ability to access the higher order category cues.

As noted previously, another interesting finding of the current study was that high-WMC individuals recalled fewer items in the random–cue condition than the random–free condition, whereas

[1] Note that there were no differences between high and low-WMC individuals in typing speed in either experiment (both ts < 1). Thus, any differences in IRTs are not likely due to differences in basic speed of processing.

Table 3. *Within- and between-category inter-response times as a function of encoding condition and WMC for Experiment 1*

	Measure			
WMC	*Random–within*	*Random–between*	*Block–within*	*Block–between*
High	2.58 (0.43)	5.60 (0.72)	3.18 (0.23)	6.38 (1.11)
Low	3.91 (0.43)	7.29 (0.72)	3.56 (0.23)	8.84 (1.11)

Note: Values in parentheses represent one standard error of the mean. WMC = working memory capacity. Inter-response times in seconds.

low-WMC individuals recalled a similar number of items in both conditions. Prior work has found a similar result whereby cued recall is actually worse than free recall when the cues presented at retrieval do not match the cues presented at encoding, and thus there is a lack of encoding specificity (Dong, 1972; Tulving, 1983; Unsworth, 2009b). This is especially true if participants expect to be tested via free recall rather than cued recall (Craik, Byrd, & Swanson, 1987; Jacoby, 1973; Unsworth, 2009b). The fact that only high-WMC individuals demonstrated worse performance on cued recall than on free recall suggests the intriguing possibility that high-WMC individuals are more vulnerable to mismatches between encoding and retrieval contexts than low-WMC individuals. That is, given that high-WMC individuals expected a free recall test, they may have engaged in strategies at encoding to link the words together, and these strategies were actually detrimental to their performance when presented with cues at recall (e.g., Jacoby, 1973; see also Cokely et al., 2006). Recent work in our laboratory has found evidence consistent with these results suggesting that high-WMC individuals are more sensitive to manipulations of encoding specificity than are low-WMC individuals (Unsworth, Brewer, & Spillers, 2011). Future work is needed to better examine these issues.

EXPERIMENT 2

The results from Experiment 1 suggested that once cues were provided to participants, WMC differences were reduced. Importantly, although the differences in recall were reduced, they were not eliminated. In accounting for categorical recall results, Patterson (1972) suggested that the ability to retrieve items subsumed under a cue was inversely related to the number of items associated with the cue (cue overload). In terms of WMC differences in recall, this would suggest that the reason low-WMC individuals differ from high-WMC individuals under cued recall could be due to differences in cue overload (Cantor & Engle, 1993; Unsworth, 2007), in which the six items in each category in Experiment 1 may have been too many for low-WMC individuals to adequately access. To examine this notion, in Experiment 2 high- and low-WMC individuals performed the same cued recall condition as that in Experiment 1 in which category cues were provided one at a time at retrieval, and participants had to retrieve as many items as possible that matched the cue. Based on Patterson (1972), we manipulated cue overload by manipulating the number of items per category. Specifically, participants were presented with lists of 24 words from four different categories. What differed from Experiment 1 was that two of the categories had only three words, whereas the other two categories had nine words. If low-WMC individuals have problems accessing words within a category based on too much cue overload, then presenting participants with only three words from a given category should alleviate this problem, leading to equivalent recall performance for high- and low-WMC individuals. When nine words are presented, however, cue overload should still be a problem leading to recall differences between high- and low-WMC individuals in cued recall consistent with Experiment 1. In addition, one problem with the results from Experiment 1 was the fact that the lists were presented in a fixed order to all participants. Although there are principal reasons for presenting the lists in a fixed order, it is possible that this ordering influenced the results. This issue is rectified in the second experiment where the lists were presented in a random order.

Method

Participant and WMC screening

Participants were recruited from the participant pool at the University of Georgia. Individuals were selected based on a z score composite of three complex span tasks and the same overall distribution as that in Experiment 1. The same prescreening procedure as that in Experiment 1 was used. Based on the composite score, participants were 23 new high-WMC individuals (z-WMC = 0.82, SD = 0.25) and 24 new low-WMC individuals (z-WMC = –0.87, SD = 0.45).

Categorized recall procedure

In this task, participants were given two lists of 24 words each. For each list, words were from four different categories with either three or nine words per category. Words and category labels were the same as those in Experiment 1. At presentation, each word was presented alone in the centre of the screen for 2 s each. Following a 1-s interval, the next word was presented. In the random encoding condition, the 24 words were presented randomly without any category labels. In the blocked encoding condition, the words were presented such that first the category label was presented alone for 3 s followed by each of the words (presented alone for 2 s each) from that category. In both lists, two of the categories had three words, and two of the categories had nine words. At recall, participants were presented with one of the category labels and were instructed to recall those items from the list that matched the category label. Participants had 30 s to recall as many items matching the category cue as possible in any order they wished by typing the words and pressing ENTER to record their response. After 30 s, the next category cue was presented, and participants were again instructed to recall the words matching the new category label. Thus, within a 2-min recall period, participants had 30 s to recall from each of four different categories. The lists and different combinations of conditions were presented in a random order.

Results

Proportion of words recalled

The proportion of words recalled was examined with a 2 (encoding: blocked vs. random) × 2 (list length: 3 vs. 9 words) × 2 (WMC: high vs. low) ANOVA with encoding and list length conditions as within-subject factors and WMC as a between-subjects factor. Shown in Table 4 are the proportions of words recalled for each of the encoding and list length conditions as a function of WMC. As can be seen, high-WMC individuals recalled more words on average than low-WMC individuals ($M = .66$, $SE = .02$ vs. $M = .58$, $SE = .02$), $F(1, 45) = 5.64$, $MSE = .05$, $p < .05$, $\eta_p^2 = .11$. A higher proportion of words was recalled with a list length of 3 than with a list length of 9 ($M = .71$, $SE = .02$ vs. $M = .53$, $SE = .02$), $F(1, 45) = 72.49$, $MSE = .02$, $p < .01$, $\eta_p^2 = .62$. As shown in Figure 2, there was a WMC × List Length interaction, $F(1, 45) = 6.30$, $MSE = .02$, $p < .05$, $\eta_p^2 = .12$, such that high-WMC individuals recalled more words with a list length of 9 than did low-WMC individuals, $t(45) = 4.24$, $p < .01$, but there were no differences with a list length of 3, $t(45) = 0.51$, $p > .61$.[2]

[2] Note that we also examined this interaction after performing a nonlinear transformation (i.e., probit transformation) of the measurement scale to examine whether the interaction was simply due to differences in sensitivity at different points of the measurement scale. Transforming the data in this way led to qualitatively identical results such that the WMC × List Length interaction remained significant, $p < .05$, $\eta_p^2 = .13$. We further examined the lack of a WMC difference on list length 3 via a Bayesian analysis suggested by Rouder, Speckman, Sun, Morey, and Iverson (2009) in which we computed the Bayes factor examining the lack of an effect. The estimated Bayes factor suggested that the odds were 4.1:1 in favour of the null. That is, the null hypothesis was more than four times more likely than the alternative. Thus, it would seem that there are probably little to no difference between high- and low-WMC individuals in recall for list length 3 in the current study.

Table 4. *Proportion correct and inter-response times as a function of conditions and WMC for Experiment 2*

	Conditions							
	Random–3		*Random–9*		*Block–3*		*Block–9*	
WMC	*PC*	*IRT*	*PC*	*IRT*	*PC*	*IRT*	*PC*	*IRT*
High	.60 (.05)	3.81 (0.35)	.54 (.03)	4.11 (0.32)	.85 (.04)	2.37 (0.16)	.65 (.03)	3.36 (0.15)
Low	.58 (.05)	4.79 (0.35)	.42 (.03)	4.64 (0.32)	.82 (.04)	3.10 (0.16)	.51 (.03)	3.91 (0.15)

Note: Values in parentheses represent one standard error of the mean. PC = proportion correct. WMC = working memory capacity. IRT = inter-response time in seconds.

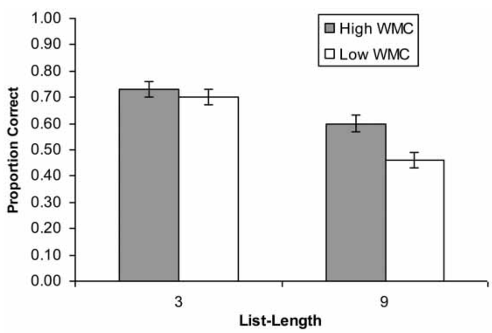

Figure 2. *Proportion correct as a function of working memory capacity and list length. Error bars represent one standard error of the mean. WMC = working memory capacity.*

Additionally, there was an effect of encoding condition such that more words were recalled with blocked presentation than with random presentation ($M = .71$, $SE = .02$ vs. $M = .54$, $SE = .02$), $F(1, 45) = 41.74$, $MSE = .03$, $p < .01$, $\eta_p^2 = .48$. This effect also interacted with list length such that the difference between list lengths was greater for blocked (M length-3 = .83, $SE = .03$ vs. M length-9 = .58, $SE = .02$) than for the random presentation (M length-3 = .59, $SE = .04$ vs. M length-9 = .48, $SE = .02$), $F(1, 45) = 9.71$, $MSE = .003$, $p < .01$, $\eta_p^2 = .18$. No other effects reached conventional levels of significance (all ps > .85).

Inter-response times

Next IRTs for items recalled for each category cue were examined with a 2 (encoding: blocked vs. random) × 2 (list length: 3 vs. 9 words) × 2 (WMC: high vs. low) ANOVA. Shown in Table 4 are corresponding IRTs. The results suggested a main effect of encoding such that IRTs were faster for blocked than for random presentation ($M = 3.19$ s, $SE = .09$ vs. $M = 4.33$ s, $SE = .20$), $F(1, 43) = 36.99$, $MSE = 1{,}612{,}075$, $p < .01$, $\eta_p^2 = .46$. There was a main effect of length such that IRTs were faster for a list length of 3 than for a list length of 9 ($M = 3.52$ s, $SE = 0.14$ vs. $M = 4.01$ s, $SE = 0.14$), $F(1, 43) = 11.31$, $MSE = 1{,}059{,}000$, $p < .01$, $\eta_p^2 = .21$. These two factors also interacted, suggesting that the difference between list lengths was greater under blocked (M length-3 = 2.74 s, $SE = 0.12$ vs. M length-9 = 3.63 s, $SE = 0.11$) than random conditions (M length-3 = 4.30 s, $SE = 0.25$ vs. M length-9 = 4.38 s, $SE = 0.23$), $F(1, 43) = 11.31$, $MSE = 176{,}315$, $p < .05$, $\eta_p^2 = .13$. Finally, there was a main effect of WMC, $F(1, 58) = 9.37$, $MSE = .001$, $p < .01$, $\eta_p^2 = .14$, suggesting that high-WMC individuals had faster IRTs than low-WMC individuals ($M = 3.41$ s, $SE = 0.17$ vs. $M = 4.11$ s, $SE = 0.17$). No other effects reached conventional levels of significance (all ps > .28).[3]

Discussion

The results from Experiment 2 suggested that high- and low-WMC individuals differed in cued recall with a long list length but there were no

[3] Note that for these IRT analyses, two low-WMC individuals were excluded due to the fact that in one condition they recalled only one item and thus did not have an IRT.

differences with a short list length. Thus, consistent with Patterson (1972), this suggests that recall is determined not only by access to higher order cues, but also by the number of items subsumed under each cue. When retrieval cues that are associated with low cue overload are provided, both high- and low-WMC individuals can access the items, leading to fairly high levels of recall. Although the WMC groups performed equivalently in the low cue overload condition, performance was not at ceiling. Thus, this suggests something like a complexity effect for individual differences in WMC such that when complexity (or self-initiated processing) in the form of cue overload is low, WMC differences do not arise, but as complexity continuously increases, WMC differences also increase. This suggests that there are not necessarily qualitative differences between high- and low-WMC individuals, rather that these differences are probably due to quantitative differences in the ability to deal with cue overload. Future work is needed to better examine these hypotheses by manipulating cue overload in a more fine-grained manner.

GENERAL DISCUSSION

In two experiments, we examined the dynamics of free and cued recall with lists of categorical words. Across both experiments, the results are consistent with the notion that covariation in WMC and recall from long-term memory is partially due to differences in accessibility. The results demonstrate that high- and low-WMC individuals differ in recall from long-term memory partly due to differences in accessibility of both higher order retrieval cues and individual words. These results are consistent with prior work examining recall deficits in children (Eysenck & Baron, 1974), older adults (Hultsch, 1975), and frontal patients (Incisa della Rocchetta & Milner, 1993), suggesting that reduced executive functioning leads to a reduced ability to strategically search long-term memory.

These results can be interpreted within a hierarchical search framework (Rundus, 1973; Shiffrin, 1970; Unsworth, 2008) in which it is assumed that first participants search for appropriate retrieval cues, and then, once a retrieval cue has been sampled, participants search for items associated to that retrieval cue. In terms of individual differences in WMC, the current results suggest that low-WMC individuals have deficits in adequately searching and accessing both the retrieval cues themselves and the items subsumed under the retrieval cues. This is clear not only from differences in the number of words recalled and the number of categories recalled, but from the IRTs analyses that demonstrated that low-WMC individuals had slower between- and within-cluster IRTs than did high-WMC individuals. Thus, low-WMC individuals cannot recall as much information from long-term memory as high-WMC individuals due to general search and access problems at both the cue and individual item level.[4]

A potential alternative explanation to the current results is that perhaps high- and low-WMC individuals differ in mental processing speed whereby low-WMC individuals are slower overall than high-WMC individuals. Given that only 30 s was given in the cued recall conditions, it possible that low-WMC individuals simply did not have enough time to recall all of the items given their lower speed of processing. Evidence consistent with this notion is that in both experiments high-WMC individuals had faster IRTs than low-WMC individuals. Thus, it is possible that differences in processing speed partially accounted for the current results. Although we did not specifically measure processing speed in

[4] One potential explanation, not considered previously, for the current results is that high- and low-WMC individuals simply differ in basic verbal ability or vocabulary size, which accounts for the differences in recall. That is, because the stimuli were words and because participants had to type their responses, it is possible that basic verbal ability differences could have influenced the results. However, as part of our initial laboratory screening procedure, all participants completed a number of cognitive ability tests including a basic vocabulary test. Examining high, and low-WMC differences in vocabulary in both experiments suggested no differences, both *t*s < 1.20, both *p*s $> .23$, and none of the results changed when partialling vocabulary out. Thus, the differences found in the current study were not due to differences in verbal abilities, per se.

the current study, we do not think that differences in processing speed accounted for the current results. Specifically, in both experiments, the time to recall the first item from memory was not different between high ($M = 3.14$, $SE = 0.81$) and low ($M = 3.24$, $SE = 0.76$) WMC individuals (both $Fs < 1$). An inspection of cumulative recall functions also suggested that both high- and low-WMC individuals reached asymptotic recall levels well before the end of the recall period. Additionally, as noted in Footnote 1, high- and low-WMC individuals did not differ in overall typing speed. Furthermore, although many studies have found a relation between WMC and processing speed in ageing studies, the relation between WMC and processing speed in younger adults is much weaker (e.g., Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Hedden, Lautenschlager, & Park, 2005; Unsworth, Spillers, & Brewer, 2011). Furthermore, recent work from our laboratory has suggested that although high- and low-WMC individuals differ in mean reaction time in a number of tasks, this difference is actually due to the fact that low-WMC individuals have proportionally more slow responses than high-WMC individuals, but there is no difference in the fastest responses (e.g., Unsworth, Redick, Lakey, & Young, 2010; see also Schmiedek, Oberauer, Wilhelm, Süß, & Wittmann, 2007). Thus, it seems unlikely that differences in processing speed account for the current results, although future work is needed to better test these hypotheses.

These results suggest that WMC differences in recall from long-term memory are due, in part, to differences in accessibility, whereby low-WMC individuals do not search as effectively as high-WMC individuals. Of course, as mentioned previously, differences in encoding abilities are also probably present and can account for some of the observed WMC differences. Furthermore, a clear unresolved issue is what exactly allows high-WMC individuals to search their memories more effectively than low-WMC individuals. That is, why are there differences in search abilities? Speculatively, we have suggested that search differences arise from differences in the overall executive search strategies and differences in the ability to self-generate retrieval cues. Yet, it is not clear how high- and low-WMC individuals differ in retrieval strategies, and it is not clear how or why they differ in the ability to self-generate retrieval cues. It is also not clear whether differences in clustering appear due to differences in the ability to self-generate cues or whether these differences are due to differences in using just recalled items to prime additional items from the list. Furthermore, it is unclear how encoding and retrieval strategies may interact differentially for high- and low-WMC individuals. Finally, although the current work suggests that WMC differences in long-term memory are partially due to accessibility issues, it should be noted that these effects are for standard list-learning paradigms, and it is not clear whether similar results will be found with more complex materials. Future work is needed to examine possible differences in strategic search and the ability to self-generate effective retrieval cues in a variety of contexts.

Original manuscript received 13 April 2011
Accepted revision received 11 April 2012
First published online 17 July 2012

REFERENCES

Bailey, H., Dunlosky, J., & Kane, M. J. (2008). Why does working memory span predict complex cognition? Testing the strategy affordance hypothesis. *Memory & Cognition*, *36*, 1383–1390.

Cantor, J., & Engle, R. W. (1993). Working memory capacity as long-term memory activation: An individual differences approach. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *19*, 1101–1114.

Cohen, B. H. (1966). Some-or-none characteristics of coding behavior. *Journal of Verbal Learning and Verbal Behavior*, *5*, 182–187.

Cokely, E. T., Kelley, C. M., & Gilchrist, A. H. (2006). Sources of individual differences in working memory: Contributions of strategy to capacity. *Psychonomic Bulletin & Review*, *13*, 991–997.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30,* 163–183.

Cowan, N., Towse, J. N., Hamilton, Z., Saults, J. S., Elliot, E. M., Lacey, J. F., et al. (2003). Children's working-memory processes: A response-timing analysis. *Journal of Experimental Psychology: General*, *132,* 113–132.

Craik, F. I., Byrd, M., & Swanson, J. M. (1987). Patterns of memory loss in three elderly samples. *Psychology and Aging*, *2,* 79–86.

Dong, T. (1972). Cued partial recall of categorized words. *Journal of Experimental Psychology*, *93*, 123–129.

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 44, pp. 145–199). New York, NY: Elsevier.

Eysenck, M. W., & Baron, C. R. (1974). Effects of cuing on recall from categorized word lists. *Developmental Psychology*, *10,* 665–666.

Healey, M. K., & Miyake, M. (2009). The role of attention during retrieval in working memory span: A dual-task study. *Quarterly Journal of Experimental Psychology*, *62,* 733–745.

Hedden, T., Lautenschlager, G., & Park, D. C. (2005). Contributions of processing ability and knowledge to verbal memory tasks across the adult lifespan. *Quarterly Journal of Experimental Psychology*, *58,* 169–190.

Hultsch, D. F. (1975). Adult age differences in retrieval: Trace-dependent and cue-dependent forgetting. *Developmental Psychology*, *11,* 197–201.

Incisa Della Rocchetta, A., & Milner, B. (1993). Strategic search and retrieval inhibition: The role of the frontal lobes. *Neuropsychologia*, *31,* 503–524.

Jacoby, L. L. (1973). Test appropriate strategies in retention of categorized lists. *Journal of Verbal Learning and Verbal Behavior*, *12,* 675–682.

Kane, M. J., & Engle, R. W. (2000). Working memory capacity, proactive interference, and divided attention: Limits on long-term retrieval. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *26,* 333–358.

Mandler, G. (1975). Memory storage and retrieval: Some limits on the reach of attention and consciousness. In P. RabbittM. A. & S. Dornic (Eds.), *Attention and Performance V*. New York, NY: Academic Press.

Mandler, G. (2011). From association to organization. *Current Directions in Psychological Science*, *20,* 232–235.

Miyake, A., & Shah, P (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. New York, NY: Cambridge University Press.

Mogle, J. A., Lovett, B. J., Stawski, R. S., & Sliwinski, M. J. (2008). What's so special about working memory? An examination of the relationship among working memory, secondary memory, and fluid intelligence. *Psychological Science*, *19,* 1071–1077.

Murdock, B. B. (1976). Item and order information in short-term serial memory. *Journal of Experimental Psychology: General*, *105,* 191–216.

Nelson, D. L., & Goodmon, L. B. (2003). Disrupting attention: The need for retrieval cues in working memory theories. *Memory & Cognition*, *31,* 65–76.

Patterson, K. E. (1972). Some characteristics of retrieval limitations from long-term memory. *Journal of Verbal Learning and Verbal Behavior*, *11,* 685–691.

Patterson, K. E., Meltzer, R. H., & Mandler, G. (1971). Inter-response times in categorized free recall. *Journal of Verbal Learning and Verbal Behavior, 10*, 417–426.

Pollio, H. R., Richards, S., & Lucas, R. (1969). Temporal properties of category recall. *Journal of Verbal Learning and Verbal Behavior*, *8,* 529–536.

Raaijmakers, J.G.? W., & Shiffrin, R. M. (1980). SAM: A theory of probabilistic search of associative memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 14). New York, NY: Academic Press.

Radvansky, G. A., & Copeland, D. E. (2006). Memory retrieval and interference: Working memory issues. *Journal of Memory and Language*, *55,* 33–46.

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., et al. (in press). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment*.

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General*, *126,* 211–227.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16,* 225–237.

Rundus, D. (1973). Negative effects of using list items as recall cues. *Journal of Verbal Learning & Verbal Behavior*, *12,* 43–50.

Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H. M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General*, *136,* 414–429.

Shiffrin, R. M. (1970). Memory search. In D.A. Norman (Ed.), *Models of human memory* (pp. 375–447). New York, NY: Academic Press.

Shiffrin, R. M., & Atkinson, R. C. (1969). Storage and retrieval processes in long-term memory. *Psychological Review, 79*, 179–193.

Tulving, E. (1983). *Elements of episodic memory*. New York, NY: Oxford University Press.

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

Unsworth, N. (2007). Individual differences in working memory capacity and episodic retrieval: Examining the dynamics of delayed and continuous distractor free recall. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *33,* 1020–1034.

Unsworth, N. (2008). Exploring the retrieval dynamics of delayed and final free recall: Further evidence for temporal-contextual search. *Journal of Memory and Language*, *59,* 223–236.

Unsworth, N. (2009a). Examining variation in working memory capacity and retrieval in cued recall. *Memory*, *17,* 386–396.

Unsworth, N. (2009b). Individual differences in self-initiated processing at encoding and retrieval: A latent variable analysis. *Quarterly Journal of Experimental Psychology*, *62,* 257–266.

Unsworth, N. (2009c). Variation in working memory capacity, fluid intelligence, and episodic recall: A latent variable examination of differences in the dynamics of free recall. *Memory & Cognition*, *37,* 837–849.

Unsworth, N. (2010). On the division of working memory and long-term memory and their relation to intelligence: A latent variable analysis. *Acta Psychologica*, *134,* 16–28.

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2009). There's more to the working memory–fluid intelligence relationship than just secondary memory. *Psychonomic Bulletin & Review*, *16,* 931–937.

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2011). Variation in working memory capacity and episodic memory: Examining the importance of encoding specificity. *Psychonomic Bulletin & Review*, *18,* 1113–1118.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114,* 104–132.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37,* 498–505.

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent variable analysis of the relationship between processing and storage. *Memory*, *17,* 635–654.

Unsworth, N., Redick, T. S., Lakey, C. E., & Young, D. L. (2010). Lapses in sustained attention and their relation to executive and fluid abilities: An individual differences investigation. *Intelligence*, *38,* 111–122.

Unsworth, N., & Spillers, G. J. (2010). Variation in working memory capacity and episodic recall: The contributions of strategic encoding and contextual retrieval. *Psychonomic Bulletin & Review*, *17,* 200–205.

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2011). Variation in verbal fluency: A latent variable analysis of switching, clustering, and overall performance. *Quarterly Journal of Experimental Psychology*, *64,* 447–466.

Watkins, M. J. (1979). Engrams as cuegrams and forgetting as cue overload: A cueing approach to the structure of memory. In C. R. Puff (Ed.), *Memory organization and structure* (pp. 173–195). New York, NY: Academic Press.

Wixted, J. T., & Rohrer, D. (1994). Analyzing the dynamics of free recall: An integrative review of the empirical literature. *Psychonomic Bulletin & Review*, *1,* 89–106.