Journal of Experimental Psychology:
Learning, Memory, and Cognition
2011, Vol. 37, No. 3, 766–775

0278-7393/11/$12.00 DOI: 10.1037/a0022475

RESEARCH REPORT

Semantic Cuing and the Scale Insensitivity of Recency and Contiguity

Sean M. Polyn
Vanderbilt University

Gennady Erlikhman
University of California, Los Angeles

Michael J. Kahana
University of Pennsylvania

In recalling a set of previously experienced events, people exhibit striking effects of recency, contiguity, and similarity: Recent items tend to be recalled best and first, and items that were studied in neighboring positions or that are similar to one another in some other way tend to evoke one another during recall. Effects of recency and contiguity have most often been investigated in tasks that require people to recall random word lists. Similarity effects have most often been studied in tasks that require people to recall categorized word lists. Here we examine recency and contiguity effects in lists composed of items drawn from 3 distinct taxonomic categories and in which items from a given category are temporally separated from one another by items from other categories, all of which are tested for recall. We find evidence for long-term recency and for long-range contiguity, bolstering support for temporally sensitive models of memory and highlighting the importance of understanding the interaction between temporal and semantic information during memory search.

Keywords: free recall, organization theory, memory search, cued recall, category

When we think back upon our own past, we have a bevy of cues to help us sort through the multitude of memories stored up over a lifetime. For example, temporal cues may be used to remember one's recent intention after stumbling downstairs for something in the middle of the night, and semantic cues may help one recall just what those prongs at the end of a fork are called. In most cases, both semantic and temporal cues are brought to bear to retrieve desired information, as when one tries to remember which vegetables one purchased on a recent trip to the supermarket. In the current study, we used the free recall paradigm to explore how temporal and semantic cues interact while people searched through their memory for recently learned information.

Theories of memory search suggest that these cues are representations that are active in the cognitive system when an experience takes place and that the components of these representations are associated with the features of the experience (e.g., Bower, 1972). To the extent that the same set of cues is present at a later point, the experience will be more accessible (e.g., Tulving & Osler, 1968); the associations between the global state of mental context and the stored memory at the time of the retrieval attempt determine the likelihood that the memory will be retrieved (Tulving, 1983). In this way, mental context acts as a retrieval cue, and researchers have characterized the ways in which temporal information (e.g., Kahana, 1996; Kahana, Howard, & Polyn, 2008), semantic information (e.g., Bousfield, 1953; Howard & Kahana, 2002b), and source information (e.g., Hintzman, Block, & Inskeep, 1972; Polyn, Norman, & Kahana, 2009) can influence memory search.

Early work on organization and memory focused on semantic organization, the tendency for items with similar meanings to be reported successively during the recall process (Bousfield, 1953; Bousfield & Sedgewick, 1944; Cofer, Bruce, & Reicher, 1966). This work used the free recall paradigm to show both that semantically related items are generally easier to remember than unrelated items (Tulving & Pearlstone, 1966) and that successive recalls made by a participant tend to come from the same category (Bousfield, 1953). More recently, theorists focused their attention on temporal organization, the tendency for items studied nearby in time to be recalled successively during memory search (Howard & Kahana, 1999; Kahana, 1996; Kahana et al., 2008). Although both semantic and temporal factors exert substantial influence on recall sequences, only a few studies have looked systematically at how they interact during memory search (e.g., Batchelder & Riefer, 1980; Borges & Mandler, 1972; Howard & Kahana, 2002b; Polyn et al., 2009).

This article was published Online First February 7, 2011.

Sean M. Polyn, Department of Psychology, Vanderbilt University; Gennady Erlikhman, Department of Psychology, University of California, Los Angeles; Michael J. Kahana, Department of Psychology, University of Pennsylvania.

This research was supported by National Institutes of Health Grant MH055687. The data reported here were collected as part of Gennady Erlikhman's senior thesis at the University of Pennsylvania, under the supervision of Sean M. Polyn and Michael J. Kahana. We thank Jonathan Miller, Neal Morton, and members of Michael J. Kahana's Computational Memory Lab for assistance with data collection.

Correspondence concerning this article should be addressed to Sean M. Polyn, Vanderbilt University, Department of Psychology, PMB 407817, 2301 Vanderbilt Place, Nashville, TN 37204. E-mail: sean.polyn@vanderbilt.edu

Howard and Kahana (2002b) examined the influence of semantic information on recall dynamics in delayed and continual distraction free recall, when the study lists were composed of items randomly drawn from a large word pool with relatively few strong semantic associates. They found that despite the random character of the lists, words that were more semantically related to one another (measured with latent semantic analysis [cos θ]; Landauer & Dumais, 1997) were more likely to be recalled successively during memory search (the semantic proximity effect). Howard and Kahana found that this effect was present both in the presence and the absence of substantial math distraction but that it was especially pronounced for items that were both semantically related and temporally proximal. This finding raises a number of questions regarding the characteristics of this interaction between temporal and semantic information, such as how it falls off with temporal distance and how it exhibits itself with increased semantic relatedness between the studied items.

A classic study by Watkins and Peynircioğlu (1983) suggests a complementary avenue of investigation, regarding the effect of strong semantic associations on the recency effect of free recall. The law of recency describes perhaps the most prominent behavioral effect in the domain of memory search: the finding that, all other things being equal, recent events are remembered better than more distant ones (T. Brown, 1824; Calkins, 1896). Watkins and Peynircioğlu, in examining the relationship of source (encoding task) context to the recency effect, devised a paradigm in which participants studied items drawn from three very distinct task contexts, such that neighboring study items were associated with different tasks. After they studied the list, participants engaged in a series of three recall periods. In each of these, the participants were asked to recall items from just one of the task contexts. Watkins and Peynircioğlu found that in the later recall periods, although participants had already engaged in recall of items from one of the other contexts (presumably a quite distracting mental operation), there was a robust and persistent recency effect. This result was surprising, as studies of the delayed free recall paradigm (Glanzer & Cunitz, 1966; Postman & Phillips, 1965) suggested that the first recall period should have disrupted the recency effect for the later recall periods. Watkins and Peynircioğlu did not examine the organization of responses during the recall periods (although later work by Howard and Kahana, 1999, showed that temporal organization of items on a study list is robust to a substantial amount of interitem distraction).

It is an open question as to whether a strong semantic cue could similarly bridge the temporal intervals associated with this paradigm and allow one to observe a persistent recency effect in the later recall periods. Watkins and Peynircioğlu (1983) suggested that this is not the case; an unpublished study in which participants studied lists composed of words from distinct taxonomic categories and then engaged in a series of recall-by-category periods is mentioned in the introduction to their paper. They failed to observe category-specific recency (for the later recall periods). Watkins and Peynircioğlu suggested that this failure stems in part from the fact that the different items, although they came from quite distinct categories, were all words, and as such shared a general similarity.

In this study, we present positive evidence that when participants study lists of items drawn from several distinct taxonomic categories, the strong semantic relations among the categorized words allow the memory system to bridge the temporal gaps separating the categorized words from one another, resulting in both a persistent recency effect and a long-range contiguity effect. The modest size of the persistent recency effect, however, may be due to the concerns raised by Watkins and Peynircioğlu (1983). We take these results as providing further support for theories of human memory in which there is continuity between short-term and long-term processes during memory search. However, these results also point toward the importance of solving the problem of how the memory system engages in selective search of memory by features other than time and of how multiple informational dimensions interact during the search process.

A few computational models of human memory have been applied to this issue recently. One approach, the context maintenance and retrieval (CMR) model, suggests that temporal, semantic, and source information combine in an internal context representation that is used to guide memory search (Polyn et al., 2009). By including multiple contextual cues, the model can explain how strong semantic associations enhance the effects of temporal recency and temporal contiguity. A second approach, the scale-independent memory, perception, and learning (SIMPLE) model, allows multiple informational dimensions to determine the location of item representations in a multidimensional space, and proximity of items in that space determines whether they will be recalled (G. D. A. Brown, Neath, & Chater, 2007; Surprenant, Neath, & Brown, 2006). Whether these models can account for the full pattern of results reported here is a question we return to in the Discussion.

Method

Participants

Twenty-three participants, age 18–30 years, each performed three sessions of the experiment. These participants were tested in accordance with University of Pennsylvania Internal Review Board guidelines and were paid $15 for each session.

Materials and List Creation

We used two word pools to create the study lists in the experiment. Words drawn from the Toronto Noun Pool (Friendly, Franklin, Hoffman, & Rubin, 1982) were used in practice trials. These trials appeared in the first session of the experiment in order to familiarize participants with the free recall task. Categorized lists were drawn from a separate word pool composed of words from a number of distinct taxonomic categories. Forty-nine categories were chosen from the word pools developed by Battig and Montague (1969) and Van Overschelde, Rawson, and Dunlosky (2003). These category norms were developed by asking participants to freely name category exemplars when given a particular category name. We excluded those category exemplars that were generated by over 50% of the participants in these norming experiments. These prototypical exemplars were excluded to discourage participants from using a generate/recognize strategy during recall. Finally, we excluded category exemplars that were highly unusual (those produced by three or fewer people in the norming experiments) and several words that could plausibly belong to more than one of the taxonomic categories used in this study (e.g., "Brown" could belong to color names or university names). A handful of categories were omitted from the experiment (e.g., weapons) to avoid use of potentially highly affective stimuli.

These categorized words were used on three types of trials. One set of trials (uncategorized free recall) contained one item from each of 24 different categories. Participants were asked to freely recall these items (verbal free recall, recorded by computer microphone). A second set of trials (categorized free recall) contained eight items from each of three categories (24 items total). Participants were asked to freely recall these items. The category exemplars were interleaved in these lists: Each set of three serial positions contained an item from each category, in pseudorandom order (subject to the restriction that items from the same category could not appear in neighboring serial positions). The third set of trials (recall-by-category) was constructed identically to the categorized free recall trials, but there were three recall periods. In each of these periods the participants were cued to recall items from only one of the three categories, but they were otherwise free to recall the items in any order. These three sets of trials were intermixed across the three sessions, as described below.

Procedure

The first session of the experiment was designed to familiarize participants with the immediate free recall paradigm and the recall-by-category variant. Participants performed six trials of immediate free recall with words drawn from the Toronto Noun Pool. This was followed by six practice trials using the categorized word pool, with two trials from each of the three experimental conditions (i.e., two uncategorized trials, two categorized free recall trials, and two recall-by-category trials). Trials from the second and third sessions were drawn from these three experimental conditions. In each session, participants performed 14 trials: four trials of uncategorized free recall, five trials of categorized free recall, and five trials of recall-by-category. Trials from the three categories were intermixed in a pseudorandom order.

On a given trial, each item appeared onscreen for 2.2 s. As soon as one item left the screen, the next was presented. After the final study item left the screen, a recall cue was presented, indicating which recall test would be required for the trial (free recall or recall-by-category). In the case of recall-by-category, a category name appeared on the screen telling the participant which items to recall. The free recall period lasted 120 s. Each of the three recall-by-category periods lasted 40 s.

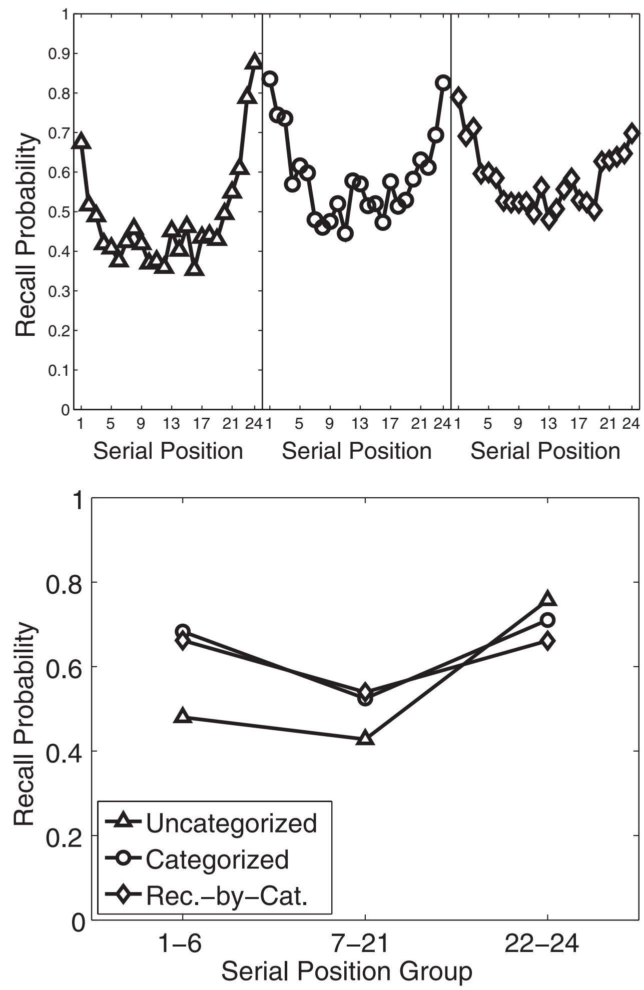

Figure 1. Serial position curves for the uncategorized free recall condition (triangles), categorized free recall condition (circles), and recall-by-category condition (diamonds). Top: Probability of recall by serial position. Bottom: Probability of recalling an item for different groups of serial positions: primacy (1–6), midlist (7–21), and recency (22–24) positions. Rec.-by-Cat. = recall-by-category.

Results

Figure 1 (top) shows the probability of recall by serial position for the uncategorized free recall (left panel), categorized free recall (middle panel), and recall-by-category (right panel) trials. We observed substantial primacy and recency effects in each of the three conditions. The presence of large primacy effects is often thought to reflect the act of rehearsal; this possibility receives more attention in the Discussion section. A number of differences can be seen in probability of recall by serial position across the three conditions; to ease comparisons between the conditions, we plotted all three conditions on the same axes (Figure 1, bottom) and present the mean proportion of items recalled for each of three sets of serial positions groups: primacy [1–6], midlist [7–21], and recency [22–24] positions (the serial position boundaries between the groups were chosen to illustrate the differences between the three curves).

Over the whole list, participants recalled significantly fewer items in the uncategorized free recall trials ($M = 0.48$, $SE = 0.02$) than in the categorized free recall trials, $M = 0.59$, $SE = 0.01$, two-tailed $t(22) = -7.19$, $p < .001$. Overall recall did not differ between the categorized free recall trials and the recall-by-category trials, $M = 0.59$, $SE = 0.02$, two-tailed $t(22) = 0.26$, $p >$ .05. A repeated-measures analysis of variance examined the effect of study condition and serial position group on recall probability (treating subject identity as a random factor). This revealed a main effect of study condition, $F(2, 44) = 22.3$, $MSE = 0.14$, $p < .001$; a main effect of serial position group, $F(2, 44) = 41.6$, $MSE = 0.78$, $p < .001$; and an interaction between the two, $F(4, 88) = 22.42$, $MSE = 0.11$, $p < .001$. Contrasts (t tests) between the categorized free recall and recall-by-category trials were not significant for any of the three serial position groups, so we combined these two conditions for a follow-up comparison to the uncategorized free recall trials. This comparison revealed significant differences between the two categorized conditions and the uncategorized free recall trials for each serial position group. For the primacy and midlist positions, performance was better on the categorized trials: primacy, $t(22) = 9.49$, $p < .001$; midlist, $t(22) = 4.75$, $p < .001$. However, for the recency positions, participants recalled significantly more items on the uncategorized trials, $t(22) = -3.42$, $p < .005$. Dual-store theories of memory (e.g., Glanzer, 1972) have taken such dissociations between early and later list positions as evidence for a short-term buffer that is responsible for recall from the terminal positions (and is insensitive to variables such as category identity), but below we discuss how this is also consistent with context-based theories.

Although the right panel of the top of Figure 1 shows the probability of recall by serial position for the recall-by-category trials, this analysis misses the most interesting dynamics of that condition, in which participants are cued to recall the items from each studied category in turn, in a series of recall periods. Figure 2 shows the probability of recall for each of these three recall-by-category periods. Here, items are plotted according to their within-category serial position on the study list. As a measure of recency, the slope of recall probability was calculated across the last five within-category serial positions (so as to span recency and midlist items); slope values that were significantly greater than zero were taken as evidence for recency. We observed a strong recency effect for the first two recall periods, which faded by the third recall period: recall period 1, mean slope = 0.054, $SE = 0.014$, $t(22) = 4.07$, $p < .001$; recall period 2, mean slope = 0.042, $SE = 0.012$, $t(22) = 3.67$, $p < .005$; recall period 3, mean slope = 0.004, $SE = 0.012$, $t(22) = 0.40$, $p > .5$. The recency effect in the second recall period appears even though this recall period started 40 s after the end of the study list. In order to rule out the possibility that this persistent recency effect was carried by trials in which no (or very few) items were recalled in the first recall period, a follow-up analysis examined the recall sequence produced in recall period 2 in more detail. When we restricted the analysis to those trials on which the participant recalled at least five items during the first recall period, the positive recency slope was still evident: mean slope = 0.024, $SE = 0.009$, $t(22) = 2.66$, $p < .05$.

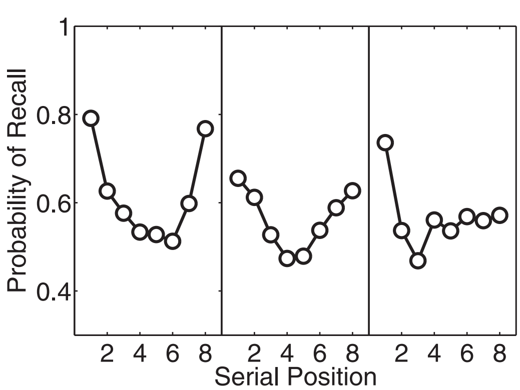

Figure 2. Within-category serial position curves for the recall-by-category trials. From left to right, each panel depicts the probability of recalling the eight studied items belonging to the cued category from the first, second, and third recall period, respectively.

The observation of the recency effect during recall-by-category suggests that the recall cue is not purely semantic, rather, a blend of temporal and semantic information is being used to search memory. However, a closer examination of temporal organization during recall-by-category suggests that temporal organization, although reliably observed, is relatively weak. Figure 3 (right panel) shows a lag–conditional response probability (lag-CRP) analysis for the responses made during recall-by-category trials. A lag-CRP analysis examines the likelihood that, given recall of an item from position N on the list, the next recalled item comes from position $N + 1$, $N - 1$, $N + 2$, $N - 2$, and so on, conditional on the availability of that item for recall. A strong contiguity effect exhibits itself as a curve that peaks for small values of lag and falls off as temporal distance increases.[1] This analysis examines only within-category transitions. Thus, items N and $N + 1$ are not neighboring items; rather, item N is an item from a particular category (say category A), and item $N + 1$ is the closest list item from the same category, in the forward direction (as such, items N and $N + 1$ can have between one and four other-category items intervening between them, owing to the pseudorandom list construction). We quantify the magnitude of the contiguity effect by calculating the *temporal factor* of the recalls made by each participant (Polyn et al., 2009). The temporal factor reflects the degree of temporal organization in a set of recalls; perfect temporal organization of recalled items results in a temporal factor of 1, and randomly organized recalled items yields an expected value near 0.5.[2] For each recalled item, we calculate the ordinal lag between it and the next item in the recall sequence and compare this observed lag to the distribution of possible lags (to the set of not-yet-recalled items). We generate a percentile rank reflecting the proportion of possible lags that correspond to more distant transitions than the observed lag and calculate the mean percentile rank across all recall transitions made by a participant in a particular condition.[3] Table 1 presents the temporal factors for each of the recall-by-category periods, as well as averaged across all three. Statistical significance was determined by comparing the observed temporal factor to a distribution of temporal factor values constructed by permuting the sequence of recalled items within each recall-by-category period. We found that the recall transitions from each recall-by-category period showed evidence for temporal organization.[4] An analysis of variance on the temporal factor by

[1] Peak values for small lags represent temporal organization, whereas peak values for distant lags represent transitions to items from primacy or recency serial positions.

[2] Inhomogeneities in the serial position curve (e.g., primacy and recency effects) cause the expected value of the temporal factor to fall below 0.5. As such, a permutation test is used to establish a baseline distribution of temporal factor scores, and observed values are compared to this distribution to determine statistical significance.

[3] For both within-category and between-category transitions, we calculate lag relative to the set of studied items that would be considered valid for a transition of that type. In other words, two items i and j from category A would have a within-category lag of 1 as long as there were no category A items intervening between them.

[4] The contiguity effect in the recall-by-category condition was somewhat dependent on transitions among early and later list items, as this effect was not significant for the second and third periods ($p > 0.1$) when serial positions 1–3 and 22–24 were excluded from the analysis.

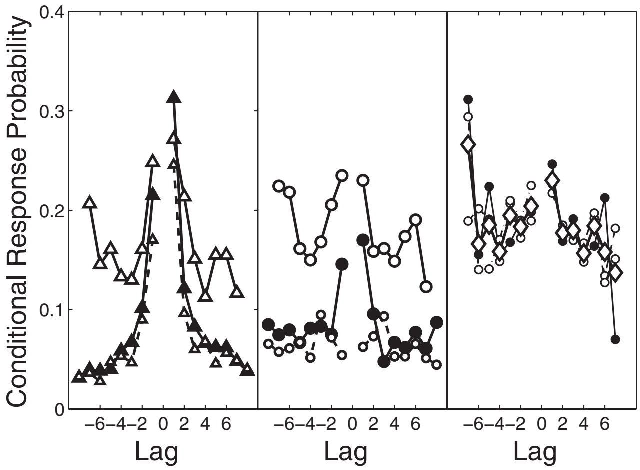

Figure 3. Lag–conditional response probability (lag-CRP) curves for the three recall conditions. Each point corresponds to the probability of making a transition of this serial position lag conditional on the availability of that item for recall. Left panel, uncategorized free recall: The solid line with unfilled triangles shows within-category transitions, and the solid line with filled triangles shows between-category transitions. For these trials, we relabeled the serial positions of the uncategorized lists with category identities drawn from the categorized lists, in order to establish a baseline for recall behavior in the absence of strong categorical cues. The dashed line corresponds to all recall transitions. Central panel, categorized free recall: The solid line with unfilled circles shows within-category transitions, and the solid line with filled circles shows between-category transitions. The dashed line corresponds to all recall transitions. Right panel, recall-by-category: The solid line with diamonds is the mean lag-CRP curve across the three recall periods. The three recall-by-category periods are respectively indicated by a solid line with filled circles, a dashed line with unfilled circles, and a dash-dotted line with unfilled circles.

recall period (treating participant identity as a random factor) found no evidence for a change in the size of the contiguity effect with recall period ($p > .4$).

In the recall-by-category trials, participants were asked to focus their memory search to items of a particular category. Between-category recall transitions (i.e., when two items from different categories were recalled successively) occurred only when participants made an error and as such were exceedingly rare. However, in the categorized free recall trials participants recalled the items from the three categories in any order, making both within-category (i.e., when two items from the same category were recalled successively) and between-category recall transitions. The presence of both within- and between-category recall transitions allows us to examine temporal organization in both the presence and the absence of a strong category association. An examination of the lag-CRP for all transitions in the categorized free recall trials (center panel, dashed line and open circles) suggests a complete absence of the contiguity effect. However, the temporal factor analysis suggests that there is weak evidence for temporal organization (see Table 1, categorized free recall), motivating a closer look. Although the lag-CRP curve for all transitions looks relatively flat, this obscures two temporal contiguity effects: When one examines within-category and between-category transitions separately, one observes a reliable contiguity effect for each class of transitions (see Table 1, Categorized free recall, Within-category, and Between-category). These dual temporal contiguity effects are obscured in the overall analysis because of the structure of the study list and the presence of strong semantic organization. On the categorized free recall trials, within-category transitions were predominant; on average participants made 65.1% within-category transitions. Because items from the same category were spaced apart on the study list, the contiguity analysis will be biased toward a null contiguity effect when category identity is not taken into account. The presence of temporal organization in the set of within-category recall transitions suggests that participants used a blend of temporal and semantic cues to search through memory.

These dual contiguity effects are depicted in Figure 3 (middle panel), which shows the lag-CRP analysis for the within-category (open circles, solid line) and between-category (filled circles, solid line) on the categorized free recall trials. The overall change in the height of the curve between the two transition types is not very important, as this is influenced by the fact that there are twice as many possible between-category transitions as there are within-category transitions (as there are three categories in the list). Of importance is the fact that both within- and between-category transitions show evidence for temporal organization. In a sense, the finding of temporal contiguity in the within-category transitions represents three simultaneous contiguity effects, one for each of the three interspersed categories. When the participant recalls an item from a particular category, the next recalled item tends to be an item both from the same category and from a nearby list position. When the participant makes a recall transition to a nearby same-category item, the recall transition must jump over items from the other two categories, as if the set of items competing for retrieval has been temporarily restricted to include just items from the most recently recalled category. The within-category transitions exhibit a temporal factor of 0.53, and the between-category transitions exhibit a temporal factor of 0.54. Each of these is significant relative to a distribution of temporal factors calculated based on permuted recall sequences (see Table 1), but the two values are not significantly different from one another, by a paired-sample t test, $t(22) = -0.58$, $p > .5$.

As we mentioned, the contiguity effect is quite weak for the recall-by-category and the categorized free recall sequences. For comparison, we examined the recall sequences produced while

Table 1
Magnitude of the Contiguity Effect for Different Recall Conditions

Recall condition	*M*	*SE*
Uncategorized free recall	0.655*	0.014
Relabeled within-category	0.563*	0.019
Relabeled between-category	0.680*	0.013
Categorized free recall	0.512*	0.009
Within-category	0.525*	0.010
Between-category	0.536*	0.016
Recall-by-category	0.518*	0.009
Recall-by-category, period 1	0.510*	0.014
Recall-by-category, period 2	0.513*	0.015
Recall-by-category, period 3	0.533*	0.015

Note. Magnitude of the contiguity effect is quantified with the temporal factor metric, a percentile-based measurement of temporal organization in free recall. * $p < .05$, by a permutation test.

participants attempted to recall words during the uncategorized free recall trials. We relabeled the items from the uncategorized condition as if they had come from a categorized list. Then these relabeled recall transitions were divided into within-category relabeled (left panel of Figure 3, unfilled triangles, solid line) and between-category relabeled (filled triangles, solid line) transitions. The contiguity effect for the full set of recall transitions is quite strong as well (unfilled triangles, dashed line). It is important to acknowledge that there are a number of differences in recall dynamics between the uncategorized and the categorized free recall conditions. For example, on the relabeled trials, within-category transitions were rare, on average participants made 23.2% within-category transitions; the threefold difference in within-category transition probability between the relabeled uncategorized and the categorized free recall trials was significant by a paired-samples t test, $t(22) = 15.9$, $p > .001$. However, despite these differences, the relabeling procedure allows us to control for the effect of serial position (and relative contiguity of same-category items) on the likelihood of making a within-category recall transition of a particular lag.

We carried out a two-way repeated-measures analysis of variance examining the effects of study condition (categorized items or uncategorized items) and category relationship (within-category or between-category) on the temporal factor (with participant identity as a random factor). We found significant main effects of both study condition, $F(1, 22) = 37.6$, $MSE = 0.19$, $p < .001$, and category relationship, $F(1, 22) = 34.2$, $MSE = 0.09$, $p < .001$, and we found that there was a significant interaction between the two factors, $F(1, 22) = 23.1$, $MSE = 0.07$, $p < .001$.

Comparing within-category transitions on relabeled uncategorized and categorized free recall trials, we found that the size of the contiguity effect was preserved. There was a marginally significant decline from 0.56 to 0.53, paired-samples t test, $t\,(22) = 1.89$, $p =$ .07. Comparing between-category transitions on relabeled uncategorized and categorized free recall trials, we found that the contiguity effect was significantly diminished (from 0.68 to 0.54), paired-samples t test, $t(22) = 8.32$, $p < .05$. Comparing within-category transitions on relabeled uncategorized free recall and recall-by-category trials, we found a significant decrease in the magnitude of the contiguity effect (from 0.68 to 0.52), paired-samples t test, $t(22) = 2.35$, $p < .05$.

We next compared the degree of temporal organization observed in the categorized free recall and recall-by-category conditions. In each of these conditions, participants used both temporal and semantic cues to target studied items on the most recent list. Although the magnitude of the contiguity effect in the recall-by-category condition was slightly lower than in categorized free recall (0.53 vs. 0.52), this difference was not significant, paired-samples t test, $t(22) = 0.646$, $p > .04$. However, participants did exhibit differential use of temporal and semantic cues in the pattern of errors that they made across these conditions. We examined two classes of recall errors. A prior-list intrusion is made when a participant recalls a word that has been studied on a previous trial (regardless of whether the trial was in a previous experimental session). On average, participants made 0.30 ($SE =$ 0.10) prior-list intrusions in the categorized free recall condition and 0.83 ($SE = 0.25$) prior-list intrusions in the recall-by-category condition.[5] A prototypical exemplar intrusion is made when a participant recalls a word that was not presented in the experiment but was one of the words excluded from the category word pool because more than 50% of the participants in the norming experiment produced this word when prompted with the category name. On average, participants made 1.26 ($SE = 0.35$) prototypical exemplar intrusions in the categorized free recall condition and 2.00 ($SE = 0.62$) prototypical exemplar intrusions in the recall-by-category condition. We combined these two classes of intrusions and found that participants made significantly more category-based recall errors in the recall-by-category condition than in the categorized free-recall condition (Wilcoxon sign rank test, $z = -2.15$, $p < .05$; tests on each recall error type considered alone were marginally significant). We take this as evidence of a stronger reliance on semantic information (or a reduced reliance on temporal information) when participants were cued explicitly for the three categories of studied items in turn.

Discussion

In this study, participants studied lists of words with items drawn from distinct taxonomic categories. These lists were either categorized (eight words from each of three categories) or uncategorized (one word from each of 24 categories). Participants then either freely recalled the studied items or engaged in a series of three recall periods, in which each of the three categories was cued in turn (recall-by-category). These conditions allowed us to examine how the interaction of semantic information and temporal information affects probability of recall by list position (in particular the recency effect) and the organization of responses (in particular the contiguity effect) in free recall.

An analysis of probability of recall by serial position found that participants generally recalled fewer items from the uncategorized study lists (except for the final items, whose recall was slightly enhanced relative to those positions in the categorized lists) and that performance was similar for the categorized free recall and recall-by-category trials. The finding of a recall disadvantage for uncategorized items in early list positions but not in later positions puts one in the mind of Glanzer's (1972) distinction between short-term and long-term stores, in which recall of recent items is due to readout from a buffer that is not sensitive to category information. However, this result is also quite consistent with a purely cue-dependent framework, in which strong semantic associations between the retrieval cue and the studied material support retrieval of items from widespread list positions and are likely to overwhelm temporal cues supporting recall of the recent items. Although the original implementations of the temporal context model (TCM; Howard & Kahana, 2002a; Sederberg, Howard, & Kahana, 2008) did not have a mechanism to represent semantic relatedness of items, Polyn et al. (2009) showed that semantic relations between items could be built into the associative structures of the network, allowing the CMR model to be applied to the interaction between temporal and semantic organization. CMR (like TCM) is a retrieved context framework; when an item is recalled, the retrieval cue is updated (i.e., modified). Thus, semantically driven recalls can diminish the recency effect: When a semantically driven recall transition is made early in the recall

[5] For comparison, participants made an average of 0.65 ($SE = 0.18$) prior-list intrusions on uncategorized free recall trials.

sequence, the end-of-list temporal information is partially supplanted by retrieved context related to the recalled item.

When participants performed the recall-by-category condition, a reliable recency effect was observed for both the first and the second recall periods. The finding of recency for the second set of items recalled by the participant is noteworthy, as this recall period occurs 40 s after the list ends, and the 40-s period is filled with the participant attempting to recall items from another category. This recency effect was still reliably observed when the analysis was restricted to those trials in which participants recalled at least five items during that first recall period, suggesting that this result is not due to covert rehearsal of these items during the first period. This result is consistent with one reported by Watkins and Peynircioğlu (1983), in which participants studied items associated with distinct task contexts and were then prompted to recall items associated with each task in turn. In the Watkins and Peynircioğlu study, a robust recency effect was also observed during the third recall period, whereas in the current study there was no evidence of the recency effect by the third recall period.

In the introduction to their paper, Watkins and Peynircioğlu (1983) pointed out that they had been unsuccessful in observing persistent recency when they used items from distinct taxonomic categories as the study materials. Without further work, it will be difficult to know which experimental variables are most critically involved in the observation of persistent recency. However, two points warrant further consideration: the relative magnitude of the primacy effect in these studies and differences in stimulus materials between the paradigms. In the current study we observed a substantial primacy effect in every condition (approaching the size of the recency effect for uncategorized free recall, rivaling the recency effect in categorized free recall, and exceeding the recency effect in recall-by-category; see Figure 1). It is unclear whether the primacy effect here is due to the distinctive nature of the stimulus materials (exemplars of taxonomic categories) or the presence of covert rehearsal (Rundus, 1971; Tan & Ward, 2000). However, retrieved context models of memory search (Howard & Kahana, 2002a; Polyn et al., 2009; Sederberg et al., 2008) suggest that an increased primacy effect will be especially devastating to the recency effect: The memory system retrieves start-of-list context upon recalling the primary items from the study list, and this context retrieval directly pushes out the residual end-of-list context underlying the recency effect. Any model of memory in which retrieval is a competitive process (i.e., most modern accounts) will make a similar prediction, although the magnitude of a primacy/recency trade-off will depend on the details of the model and the ferocity of the competition. In comparison to that seen in the current study, the primacy effect observed by Watkins and Peynircioğlu was quite modest (much less than half the size of the recency effect). This suggests that perhaps the deep encoding tasks used in that study disrupted any attempts by the participants to engage in active rehearsal. It is possible that by reducing the size of the primacy effect in the current experiment, one would increase the size of the persistent recency effect to levels comparable to those observed by Watkins and Peynircioğlu. One possibility would be to introduce an effortful encoding task during study, to disrupt any active rehearsal processes employed by the participants. Such a manipulation was not employed here, in order that we might examine the interaction between semantic and temporal information, without further taking into account the source context introduced by such a task.

A second important difference from the Watkins and Peynircioğlu (1983) paradigm is that the items from a particular category in the current paradigm may often have substantial semantic associations to items from other categories studied in the same list. These intercategory similarities might generally weaken the size of both the persistent recency effect and the long-range contiguity effect (in that the same retrieval cue will target words from multiple categories). In the Watkins and Peynircioğlu paradigm, the words were selected at random (or at least were not drawn from strong taxonomic categories), which may have led to less interitem interference in that study. It is also possible that the presence of strong source contexts reduced the likelihood that semantic associations between words studied in different contexts would be discovered by the participants.

Recent work by Farrell and Lewandowsky (2008) uncovered a tendency for participants to make distant forward recall transitions during the first recall transition (e.g., a recall of a midlist item followed by a recall of an end-of-list item) in a number of free recall data sets. Polyn et al. (2009) replicated this finding and interpreted it as a recency effect that persists beyond the first output position, pulling the search process back to the end of the list. Polyn et al. showed that this phenomenon is consistent with CMR (as well as TCM; Howard & Kahana, 2002a) and arises because the temporal context associated with end-of-list items is only partially disrupted when a midlist item is recalled. Further work is needed to determine whether this effect, which exhibits itself as a forward nonmonotonicity in the lag-CRP curve for the first recall transition, reflects the same mechanism as underlies the persistent recency effect in the current study.

The central analyses of this study examined the contiguity effect for the different list types (see Table 1 and Figure 3). We used the temporal factor statistic (Polyn et al., 2009) to quantify the magnitude of the contiguity effect (in essence the degree of temporal clustering) as a single number. We showed that although temporal organization looked very weak when we examined the recall sequences from the categorized free recall trials overall, robust contiguity effects were observed for these sequences when within-category and between-category transitions were considered separately. Furthermore, we found that the magnitude of the contiguity effect for within-category transitions was preserved relative to between-category transitions in the same lists, although the within-category items were widely spaced throughout the study list. This result extends the finding, reported by Howard and Kahana (1999), that the contiguity effect in free recall persists even when the studied items are spaced with the addition of interpolated distracting activities (in that case, mental arithmetic). Here, the distracting mental activities involved studying items drawn from two other categories, arguably a more distracting situation, as participants had to recall items from all three categories during the recall period. This finding suggests that when a person recalls an item from a particular category, the items from the other categories (despite their proximity to the just-recalled item and the ubiquity of the contiguity effect; Kahana et al., 2008) do not compete effectively in the recall competition to determine what item is recalled next. Furthermore, the statistical interaction of study condition and category relationship in determining the temporal factor suggests that the cognitive process underlying the recall competition itself involves an interaction between temporal and categorical information. A pure category cue, which equally supported recall of all items from a particular category, would account for the increased likelihood of making same-category recall transitions but would not account for the preserved reliance on temporal information (relative to similarly spaced items on a list with no category structure). This second point is perhaps best captured by a model in which categorical and temporal information interact during a recall competition, a point that we return to below.

In a final analysis, we presented evidence that participants relied more on semantic cues in the recall-by-category condition than in the categorized free recall condition. Participants made significantly more semantically based recall errors in the recall-by-category condition (both prior-list intrusions and prototypical exemplar intrusions). The prior-list intrusions can be considered semantically based errors, as they almost exclusively involved intruding an item that was from one of the three taxonomic categories on the current list; these categories were not repeated within session for the categorized trials, so often the prior-list intrusion would originate from a list in the previous experimental session (or from an uncategorized control list in the same session). This finding suggests that providing participants with an explicit semantic retrieval cue (the category label) changes the recall dynamics in a number of ways, resulting in diminished temporal organization (Figure 3, right panel) and an increase in semantically based intrusions.

Detailed computational modeling of these phenomena is beyond the scope of this report. Nevertheless, we believe these findings inform our broader understanding of the human memory system. Crowder (1982) argued that models of human memory that rely on a short-term buffer to explain the recency effect are challenged by the finding of long-term recency. Howard and Kahana (1999) argued, along similar lines, that short-term buffer models are challenged by the long-range contiguity effect. There are very few pure short-term buffer models of human memory; nearly all models of memory are to some extent driven by a retrieval cue. For example, the hybrid context/buffer model of Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, and Usher (2005) proposes that immediate recency is supported by a short-term buffer and long-term recency is supported by a temporal context mechanism; even the search of associative memory model of human memory proposes that once the short-term buffer is emptied, a cue-driven search commences (Atkinson & Shiffrin, 1968; Raaijmakers & Shiffrin, 1981). Thus, the utility of these findings is not that they invalidate certain classes of models but rather that they help us consider the dynamics underlying a cue-driven memory system. Two classes of temporally based models are most relevant to this discussion: temporal distinctiveness models (e.g., G. D. A. Brown et al., 2007) and temporal context models (Howard & Kahana, 2002a; Polyn et al., 2009; Sederberg et al., 2008). Each class suggests that the accessibility of items falls off as a function of time or events (Lewandowsky, Brown, Wright, & Nimmo, 2006).

G. D. A. Brown et al. (2007) presented a temporal distinctiveness model called SIMPLE. According to SIMPLE, the discriminability of events relative to one another is a major factor in determining their accessibility in memory; as events recede into the past, they crowd each other on a temporal dimension, becoming less accessible. Similarity on any representational dimension (e.g., semantic similarity or source similarity) causes items to become less discriminable from one another, by the following equation (Equation 3 from G. D. A. Brown et al., 2007):

$$P(R_i) = D_i = \frac{1}{\sum_{j=1}^{n}(\eta_{i,j})}, \tag{1}$$

in which the probability of recall, $P(R)$, of an item from serial position (i), is equivalent to its discriminability (D) and is inversely proportional to the sum of the similarity (η) of the item to all of the n potentially retrievable items in memory (including itself). This equation explains why, as one increases the phonological similarity between items in a serial recall experiment (as in Henson, Norris, Page, & Baddeley, 1996, Figure 20 in Brown et al., 2007), recall performance declines. Relative to a list in which the items are well distributed on the phonological dimension (dissimilar items), the increased similarity of the phonologically related items to one another increases the η term in the denominator of Equation 1 and thus reduces the probability of recall of these items.[6]

This distinctiveness mechanism may make it challenging for SIMPLE to explain the benefit of semantic relatedness in the free recall paradigm. For example, in the current paradigm, items on the uncategorized free recall trials are well distributed in terms of their semantic relatedness (24 items, each from a different taxonomic category), whereas items on the categorized free recall trials are considerably more crowded (eight items from each of three taxonomic categories). According to SIMPLE, this increased crowding on the semantic dimension ought to decrease recall performance, but instead the categorized items are better recalled (consistent with a number of classic results from the free recall literature; e.g., Tulving & Pearlstone, 1966).[7]

It is more difficult to work out the predictions of SIMPLE with regard to the organizational effects characterized in this study. The original version of SIMPLE (G. D. A. Brown et al., 2007) did not contain machinery to simulate the order of responses during free recall. In other words, SIMPLE cannot explain how the memory system produces a string of items from a particular category, either in the case where the participant is asked to report only items from a particular category or in the case of free recall, when items spontaneously group along dimensions of similarity (e.g., Polyn et al., 2009). Recently, G. D. A. Brown, Chater, and Neath (2008) described how SIMPLE could be extended to produce organizational effects. They suggested that when a particular item is recalled, the next recalled item is likely to be one that is near the first one in a multidimensional similarity space (in our example, the dimensions of this space correspond to the semantic and temporal characteristics of the studied items).

Retrieved context models, such as TCM (Howard & Kahana, 2002a; Sederberg et al., 2008) and CMR (Polyn et al., 2009), provide a natural mechanism for understanding both facilitation of recall by similarity and organization by similarity during free recall: the contextual retrieval cue. In the CMR model (Polyn et al.,

[6] In contrast, grouping items temporally (by manipulating the spacing between sets of items, as in Figure 22 of G. D. A. Brown et al., 2007) can result in increased recall performance for all studied items.

[7] Exploratory simulations of the categorized and uncategorized free recall trials with the SIMPLE model yielded results consistent with this discussion; however, it is possible that other variants of the SIMPLE model would be able to capture this effect (see, e.g., Neath & Brown, 2006; Figure 10).

2009) multiple memory cues (semantic, temporal, and source-related) interact during the search process. By this framework, participants can explicitly target items that match the current contextual retrieval cue. Because similar items are associated to similar states of context, the context retrieved upon recalling one item from a particular category will lead to successive recalls from the same category (both facilitating recall performance and producing semantic organization). In order to cue memory to recall items from a single category (as in the recall-by-category trials), one loads up the context representation with information related to that category, which allows items associated with the same category to win a recall competition (Polyn et al., 2009).

However, the results described by Watkins and Peynircioğlu (1983) may provide a challenge for context-based models. As mentioned, participants were presented with three successive recall periods; in each one they were asked to target items associated with one of the three task contexts. A persistent recency effect was observed in each of the three recall periods. The challenge to context-based models arises because, unlike in the continual distraction paradigm, where one performs an unrelated task during the delay periods, in this paradigm one actively engages in memory search during the earlier recall periods. For a context-based model, the act of searching through memory involves retrieving the temporal context associated with the remembered items. Once a number of items have been retrieved from throughout the list, temporal context will no longer preferentially support the end-of-list items, leading to a prediction of a null recency effect in the later recall periods. Thus, although the principles of CMR are broadly consistent with the current study, it is possible that certain phenomena (such as the persistent recency effect) will challenge the model.

References

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2, pp. 89–105). New York, NY: Academic Press.

Batchelder, W., & Riefer, D. (1980). Separation of storage and retrieval factors in free recall of clusterable pairs. *Psychological Review, 87,* 375–397.

Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph, 80*(3, Pt. 2), 1–46.

Borges, M. A., & Mandler, G. (1972). Effect of within-category spacing on free recall. *Journal of Experimental Psychology, 92,* 207–214.

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology, 49,* 229–240.

Bousfield, W. A., & Sedgewick, C. H. W. (1944). An analysis of sequences of restricted associative responses. *Journal of General Psychology, 30,* 149–165.

Bower, G. H. (1972). Stimulus-sampling theory of encoding variability. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 85–121). New York, NY: Wiley.

Brown, G. D. A., Chater, N., & Neath, I. (2008). Serial and free recall: Common effects and common mechanisms? A reply to Murdock (2008). *Psychological Review, 115,* 781–785.

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review, 114,* 539–576.

Brown, T. (1824). *Lectures on the philosophy of the human mind.* Philadelphia, PA: Grigg & Bason.

Calkins, M. W. (1896). Association: An essay analytic and experimental. *Psychological Review Monograph Supplement, 1*(2).

Cofer, C. N., Bruce, D. R., & Reicher, G. M. (1966). Clustering in free recall as a function of certain methodological variations. *Journal of Experimental Psychology, 71,* 858–866.

Crowder, R. G. (1982). The demise of short-term memory. *Acta Psychologica, 50,* 291–323.

Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J., & Usher, M. (2005). The demise of short-term memory revisited: Empirical and computational investigations of recency effects. *Psychological Review, 112,* 3–42.

Farrell, S., & Lewandowsky, S. (2008). Empirical and theoretical limits on lag recency in free recall. *Psychonomic Bulletin & Review, 15,* 1236–1250.

Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). The Toronto word pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavior Research Methods & Instrumentation, 14,* 375–399.

Glanzer, M. (1972). Storage mechanisms in recall. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 5, pp. 129–193). New York, NY: Academic Press.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5,* 351–360.

Henson, R. N. A., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 49*(A), 80–115.

Hintzman, D. L., Block, R. A., & Inskeep, N. R. (1972). Memory for mode of input. *Journal of Verbal Learning and Verbal Behavior, 11,* 741–749.

Howard, M. W., & Kahana, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 923–941.

Howard, M. W., & Kahana, M. J. (2002a). A distributed representation of temporal context. *Journal of Mathematical Psychology, 46,* 269–299.

Howard, M. W., & Kahana, M. J. (2002b). When does semantic similarity help episodic retrieval? *Journal of Memory and Language, 46,* 85–98.

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition, 24,* 103–109.

Kahana, M. J., Howard, M. W., & Polyn, S. M. (2008). Associative retrieval processes in episodic memory. In H. L. Roediger III (Series Ed.) & J. Byrne (Vol. Ed.), *Learning and memory: Vol. 2. Cognitive psychology of memory* (pp. 467–490). Oxford, England: Elsevier.

Landauer, T. K., & Dumais, S. T. (1997). Solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104,* 211–240.

Lewandowsky, S., Brown, G. D. A., Wright, T., & Nimmo, L. M. (2006). Timeless memory: Evidence against temporal distinctiveness models of short-term memory for serial order. *Journal of Memory and Language, 54,* 20–38.

Neath, I., & Brown, G. D. A. (2006). SIMPLE: Further applications of a local distinctiveness model of memory. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 201–243). San Diego, CA: Academic Press.

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review, 116,* 129–156.

Postman, L., & Phillips, L. W. (1965). Short-term temporal changes in free recall. *Quarterly Journal of Experimental Psychology, 17,* 132–138.

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

Rundus, D. (1971). An analysis of rehearsal processes in free recall. *Journal of Experimental Psychology, 89,* 63–77.

Sederberg, P. B., Howard, M. W., & Kahana, M. J. (2008). A context-based theory of recency and contiguity in free recall. *Psychological Review, 115,* 893–912.

Surprenant, A., Neath, I., & Brown, G. (2006). Modeling age-related differences in immediate memory using SIMPLE. *Journal of Memory and Language, 55,* 572–586.

Tan, L., & Ward, G. (2000). A recency-based account of the primacy effect in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1589–1626.

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

Tulving, E., & Osler, S. (1968). Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology, 77,* 593–601.

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

Van Overschelde, J., Rawson, K., & Dunlosky, J. (2003). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language, 50,* 289–335.

Watkins, M. J., & Peynircioğlu, Z. F. (1983). Three recency effects at the same time. *Journal of Verbal Learning and Behavior, 22,* 375–384.

Received April 5, 2010
Revision received November 1, 2010
Accepted November 9, 2010 ■