



Published in final edited form as:

J Exp Psychol Learn Mem Cogn. 2014 May ; 40(3): 755–764. doi:10.1037/a0035585.

A retrieved context account of spacing and repetition effects in free recall

Lynn J. Lohnas and
University of Pennsylvania

Michael J. Kahana
University of Pennsylvania

Abstract

Repeating an item in a list benefits recall performance, and this benefit increases when the repetitions are spaced apart (Madigan, 1969; Melton, 1970). Retrieved context theory incorporates two mechanisms that account for these effects: contextual variability and study-phase retrieval. Specifically, if an item presented at position i is repeated at position j , this leads to retrieval of its context from its initial presentation at i (study-phase retrieval) and this retrieved context will be used to update the current state of context (contextual variability). Here we consider predictions of a computational model that embodies retrieved context theory, the context maintenance and retrieval model (CMR; Polyn, Norman, & Kahana, 2009). CMR makes the novel prediction that subjects are more likely to successively recall items that follow a shared repeated item (e.g. $i + 1, j + 1$) because both items are associated with the context of the repeated item presented at i and j . CMR also predicts that the probability of recalling at least one of two studied items should increase with the items' spacing (Lohnas, Polyn, & Kahana, 2011). We tested these predictions in a new experiment and CMR's predictions were upheld. These findings suggest that retrieved context theory offers an integrated explanation for repetition and spacing effects in free recall tasks.

Introduction

The beneficial mnemonic effects of spaced over massed learning is one of the most widely documented and well studied memory phenomena. Although it is seen in a wide range of tasks, the spacing effect is particularly strong in free recall, where the probability of recalling a repeated word often increases monotonically to spacings of 20 or more items (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Delaney, Verhoeijen, & Spiguel, 2010; Donovan & Radosevich, 1999; Madigan, 1969; Melton, 1970).

Two major theories that have been advanced to explain the spacing effect in free recall are contextual variability theory and study-phase retrieval theory. Contextual variability theory explains the spacing effect by assuming that each presented item is associated with a slowly drifting context representation, such that items repeated at greater spacings benefit from

more distinctive retrieval cues (Bower, 1972; Estes, 1955; Melton, 1970). Study-phase retrieval theory posits that repetition of an item retrieves one's memories of the repeated item's earlier occurrences and their associated contexts (Greene, 1989; Thios & D'Agostino, 1976). This retrieved information, in turn, becomes associated with the repeated item, thus providing an additional set of retrieval cues for the repetition.

Although study-phase retrieval is frequently presented as an alternative to contextual variability theory, the two accounts are not mutually exclusive (Benjamin & Tullis, 2010; Glenberg, 1979; Raaijmakers, 2003; Verkoijen, Rikers, & Schmidt, 2004; Young & Bellezza, 1982). Indeed, retrieved context models of free recall embody both mechanisms: Items are associated with an evolving representation of temporal context, whereby the contexts retrieved from repeating (or recalling) items update the current state of context (Howard & Kahana, 2002). In a complementary way, the current state of context is used as the cue for recall.

Retrieved context models have been applied to a wide range of recall phenomena involving recall of once-presented items (Gershman, Moore, Todd, Norman, & Sederberg, 2012; Howard, 2004; Howard & Kahana, 1999, 2002; Howard, Kahana, & Wingfield, 2006; Sederberg, Gershman, Polyn, & Norman, 2011; Sederberg, Howard, & Kahana, 2008; Shankar & Howard, 2012). For instance, because these models assume that each item is associated with a slowly drifting temporal context, neighboring items' context states are more strongly correlated (Figure 1A). Thus, when an item is recalled and its context is retrieved, the state of context used to cue recall of the next item will favor neighbors of the just-recalled item. As a result, these models predict the contiguity effect, or tendency to successively recall neighboring list items (Howard & Kahana, 2002; Kahana, 1996). These models also predict the heightened recall of end-of-list items in immediate free recall, termed the recency effect, because the context at the beginning of the recall period most strongly cues items from the end of the list. In addition, such models predict that the recency effect is sharply attenuated in delayed free recall because the distractor interval shifts context away from cuing recency items.

In the present work we examine the spacing and repetition effect predictions of a particular version of retrieved context theory—the Context Maintenance and Retrieval model (CMR; Polyn, Norman, & Kahana, 2009). Going beyond the analysis of well-established repetition effects, we show that CMR makes specific predictions concerning the way item repetition will influence the organization of recall, as described below. Finally, we report a new experiment that aims to test these predictions.

Simulating repetition and spacing effects in the context maintenance and retrieval model

We first examine whether the CMR model produces the qualitative effects of repetition and spacing observed in many free recall experiments. We then turn to novel predictions of the CMR model concerning recall in lists with repeated items. Rather than fitting CMR to a particular dataset we assessed the model's predictions for standard parameter values reported in previous work (Table 1). Before reporting the simulation results we describe the

structure of the model in more detail, including the major equations that govern memory encoding and retrieval.

Model structure

According to CMR, the vector representations of items (denoted \mathbf{f}_i) and context (denoted \mathbf{c}_i) interact through associative matrices (M^{FC} and M^{CF}) that are updated according to a standard Hebbian learning rule, such that $\Delta M^{FC} = \mathbf{c}_i \mathbf{f}_i^\top$ and $\Delta M^{CF} = \mathbf{c}_i \mathbf{f}_i^\top$. The relative contribution of the updated associations to the pre-existing associations is determined by a parameter γ_{FC} . As explained in Polyn et al. (2009), the model embeds the semantic associations between items in the initial values of the context-to-item associative matrix, which are determined using Latent Semantic Analysis (LSA; Landauer & Dumais, 1997). CMR assumes a primacy gradient of attention such that the change in M^{FC} is scaled by φ_i , which is greatest for early list items (all else being equal) and decreases exponentially to an asymptotic value over the course of list presentation, $\varphi_i = \varphi_s e^{-\varphi_d(i-1)} + 1$, where φ_s and φ_d are model parameters.

When an item i is presented to the model (irrespective of whether it is being presented for

the first or second time in a list), it creates a new input to context, $\mathbf{c}_i^{\text{IN}} = \frac{M^{FC} \mathbf{f}_i}{\|M^{FC} \mathbf{f}_i\|}$. Context evolves according to the equation:

$$\mathbf{c}_i = \rho_i \mathbf{c}_{i-1} + \beta \mathbf{c}_i^{\text{IN}} \quad (1)$$

where ρ_i is defined such that $\|\mathbf{c}_i\| = 1$. Context is a weighted sum of contextual states, and recent states dominate the representation. β is a model parameter that determines how much \mathbf{c}_i changes with each studied item. The rate of context updating can differ between encoding and recall events (β_{enc} and β_{rec} , respectively). To simulate delayed free recall, we assume that \mathbf{c} further evolves according to Equation 1 as subjects engage in a distracting mental activity. In this case, we consider the distractor to be a single orthogonal item with a value of β_{dist} to determine the change in \mathbf{c} .

At the time of recall, cuing with \mathbf{c}_i retrieves a vector $\mathbf{f}^{\text{IN}} = M^{CF} \mathbf{c}_i$. To determine which item the model recalls, \mathbf{f}^{IN} serves as the input to a leaky, competitive accumulation process (Usher & McClelland, 2001) whose value at time step t is determined by

$$\begin{aligned} \mathbf{x} &= (1 - \tau\kappa - \tau\lambda\mathbf{N}) \mathbf{x}_{t-1} + \tau \mathbf{f}^{\text{IN}} + \epsilon, \\ \tilde{\epsilon} &\sim \mathcal{N}(0, \eta), \\ \mathbf{x}_t &\rightarrow \max(\mathbf{x}_t, 0). \end{aligned} \quad (2)$$

Each element of \mathbf{x}_t corresponds to an element in \mathbf{f}^{IN} . τ is a time constant, κ is a leak parameter, η is a noise parameter, and λ is a parameter that controls lateral inhibition, by scaling the strength of an inhibitory matrix \mathbf{N} which connects each accumulator to all of the others except itself.

This process runs iteratively until one of the accumulating elements crosses a threshold or until the recall period is over. When an item wins the recall competition, it is re-presented to

the model, updating context according to Equation 1. The updated state of context activates a different set of features on \mathbf{f}^N and the recall competition begins again.

Results

To evaluate spacing and repetition predictions of CMR we simulated a delayed free recall experiment involving two sets of lists: control lists of once-presented items and mixed lists which contained both once-presented and repeated items. Each list contained 40 unique positions. In the control lists a unique item occupied each position. In the mixed lists 6 pairs of repeated items were mixed among 28 once presented items. Across lists there were an equal number of items pairs repeated at spacings of $\text{lag} \in \{0, 1, 2, \dots, 8\}$, where lag is defined as the number of intervening items between an item's repetitions. We used the CMR parameter values obtained by Polyn et al. (2009) in their simulation of Murdock's classic (1962) serial position effect data. To attenuate potential effects of recency we simulated a brief arithmetic distractor task following list presentation using the β_{dist} parameter reported in Sederberg et al.'s (2008) simulation of delayed and continual distractor free recall. To obtain stable predictions we ran the model on 8,400 lists (equally divided among the two list conditions).

We first considered CMR's predictions of well-established repetition effects. CMR predicts higher recall probability for massed repetitions (recall probability = 0.76) than for once-presented items (recall probability = 0.37; Madigan, 1969; Melton, 1970). CMR also predicts the spacing and lag effects, as recall probability for spaced items increases with lag (Figure 2). For this analysis, we controlled for serial position effects that may artificially give rise to the lag effect owing to the tendency for spaced items to occupy more favorable serial positions. We generated as many trials as the simulated data by matching the recall probability of each item to the serial position curve. On a given trial, the recall probability of each item was calculated randomly and independently. In Figure 2, we show CMR's predictions of recall probability subtracting out the expected recall probability based on the randomly generated lists. Because the randomized recall probabilities are usually larger than the observed recall probabilities, for illustrative purposes we added the mean marginal probability to each value.

CMR predicts the repetition effect because it assumes that the second presentation of an item leads to the retrieval of the context from its first presentation (compare Figure 1A and 1B). In this way, a repeated item is associated with an amalgam of two context states (each of which can serve as an effective retrieval cue for that item) and a once-presented item is only associated with one context state. CMR predicts the advantage for items repeated with larger spacing because the context retrieved from an item's first presentation will be more distinct from the current context. For instance, whereas in Figure 1B the repeated item's retrieved context is the previous context state, for the spaced item in Figure 1C the retrieved context is more distinct from the current context state, and thus in the latter case the repeated item benefits from being associated with two contexts which are more distinctive from each other.

Given that the probability of recalling an item repeated in positions i and j is equivalent to the probability of recalling either the occurrence of the item in position i or the occurrence

of the item in position j , contextual variability predicts that the probability of recalling either of two once-presented items in positions i and j , termed the OR score, should increase with their spacing $|i - j|$ (Lohnas, Polyn, & Kahana, 2011; Ross & Landauer, 1978). Whereas contextual variability theory and study-phase retrieval both contribute to the spacing effect, only contextual variability contributes to this property of once-presented items. In a meta-analysis of 6 free-recall studies containing lists of once-presented items, Lohnas et al. (2011) showed that OR scores reliably increase with lag. Here we assess whether CMR predicts this result by probing CMR's predictions of OR scores in control lists. We controlled for serial position effects in an analogous way to the lag analysis. We generated recall trials from CMR's simulated data with the same distributions of recall probabilities as CMR's simulated serial position curve for control lists. Recall of each item was calculated randomly and independently across trials. We present CMR predictions of the OR scores minus the expected OR scores from the randomized lists. Figure 3 shows this difference with the mean value of the random lists added back at each lag. For control lists (open squares), CMR predicts that the OR score for successively presented items should be less than that of spaced once-presented items, because the context states of spaced items are less highly correlated (Figure 1A).

Given CMR's accurate prediction of OR scores in control lists, we next considered a novel prediction of CMR: OR scores of once-presented items in mixed lists. In contrast to control lists, the presentation of repeated items leads to retrieval of prior experimental contexts, and thus the assumption of strict contextual variability no longer applies in mixed lists. For instance, consider the list of items in the mixed list represented in Figure 1C. Because item 5 is repeated at serial position 8, this retrieves the associated context from item 5, which includes the context from item 4. As a result, the context from item 9 is more similar to the context of item 4 than of item 7. This is not the case in the presentation of once-presented items as in Figure 1A: the context of item 9 is less similar to the context of item 4 than of item 7. Thus, in general context does not strictly drift the same for each presented item (with respect to Figure 1C, context does not progress perfectly from black to white), and so the similarity in contexts between two items may not strictly decrease with lag. Nonetheless, as shown in Figure 3 (filled squares) CMR predicts that the OR score should increase with lag in mixed lists.

Lastly, we considered a novel prediction of CMR regarding neighbors of repeated items in mixed lists. According to CMR, the context associated with a repeated item includes the context associated with its earlier presentation. As such, the context after the second presentation of a repeated item is more similar to the context after the first presentation than would be expected in the case of non-repeated items. Furthermore, the context at the time of test will be more similar to both repetitions than would be the case for a list with non-repeated items (e.g., compare the context for the item repeated at position 8 in Figure 1C to the context of the corresponding item in Figure 1A). The retrieved context assumption of CMR thus increases the associative strength between the second occurrence of an item (and its neighbors) with the neighbors of the first occurrence of the item. For instance, when contrasted with the corresponding items in Figure 1A, in Figure 1C the items in serial positions 9 and 10 have contexts that are more similar to items 6 and 7, due to the item

repeated at serial position 8. Thus, if an item is repeated at positions i and j , then CMR predicts that items presented after j (i.e. $j + 1, j + 2$) should share more similar context states to items presented after i ($i + 1, i + 2$).

To test CMR's novel prediction, we considered transitions between items following a shared repeated item. We calculated the proportion of those items recalled in $S_j = \{j + 1, j + 2\}$ of which CMR then recalled an item in the set $S_i = \{i + 1, i + 2\}$. We also calculated the proportion of recalls S_i of which CMR then transitioned to an item in the set S_j . We calculated the proportion of transitions for each of lags $j - i = 4$, and Figure 4 shows the mean percent of transitions across these lags. To estimate the proportion of transitions that CMR would make at these lags in the absence of repeated items, we considered transitions in control lists matched to the same serial positions considered in the mixed lists. We matched these serial positions to 100 random shuffles of the control lists, and took the mean across the reshuffled datasets. In Figure 1, we would consider the transitions between $S_j = \{9, 10\}$; $S_i = \{6, 7\}$ in C as the actual transitions, and transitions between $S_j = \{9, 10\}$; $S_i = \{6, 7\}$ in A could be considered a sample baseline condition. CMR predicts many more transitions among items that follow repeated items than expected by their yoked controls (17% vs. 4%; Figure 4).

Experiment

We designed an experiment to assess CMR's novel predictions concerning OR scores in mixed lists and the increased percent of transitions between items that follow the same repeated item.

Methods

Across 4 sessions, 35 subjects performed delayed free recall of 48 lists. Subjects were University of Pennsylvania undergraduates, graduates and staff, age 18-32. List items were drawn from a pool of 1638 words taken from the University of South Florida free association norms (Nelson, McEvoy, & Schreiber, 2004; Steyvers, Shiffrin, & Nelson, 2004, available at http://memory.psych.upenn.edu/files/wordpools/PEERS_wordpool.zip). Within each session, words were drawn without replacement. Words could repeat across sessions so long as they did not repeat in two successive sessions. Words were also selected to ensure that no strong semantic associates co-occurred in a given list (i.e., the semantic relatedness between any two words on a given list, as determined using WAS (Steyvers et al., 2004), did not exceed a threshold value of 0.55).

Subjects encountered four different types of lists: 1) control lists that contained all once-presented items; 2) pure massed lists containing all twice-presented items; 3) pure spaced lists consisting of items presented twice at lags 1-8, where lag is defined as the number of intervening items between a repeated item's presentations; 4) mixed lists consisting of once presented, massed and spaced items. Within each session, subjects encountered three lists of each of these four types. In each list there were 40 presentation positions, such that in the control lists each position was occupied by a unique list item, and in the pure massed and pure spaced lists, 20 unique words were presented twice to occupy the 40 positions. In the mixed lists 28 once-presented and six twice-presented words occupied the 40 positions. In

the pure spaced lists, spacings of repeated items were chosen so that each of the lags 1-8 occurred with equal probability. In the mixed lists, massed repetitions (lag=0) and spaced repetitions (lags 1-8) were chosen such that each of the 9 lags of 0-8 were used exactly twice within each session. The order of presentation for the different list types was randomized within each session. For the first session, the first four lists were chosen so that each list type was presented exactly once. An experimenter sat in with the subject for these first four lists, though no subject had difficulty understanding the task.

A pilot study had us concerned that subjects may be confused as to whether repeated items should be recalled twice in the recall period. To prevent our free recall study from simultaneously being a judgment of frequency task, we notified subjects of our repetition manipulation at the beginning of the experiment. Specifically, each subject read the following text on the computer screen, and the experimenter verified that the subject understood the instruction: “Within each list, you may notice that some or all words are presented twice. In other lists, you may notice that some or all words are presented only once. You only need to say each word once during the recall period, even if it was presented more than once on the list.”

Prior to the start of each list, a 1500 ms blank screen followed by a 1000 ms fixation cross indicated the start of the list. Each word was presented on the screen for 2 s, followed by an 800 ms interstimulus interval. Subjects were instructed to focus on each word as it appeared on the screen, keeping in mind that they would later attempt to recall as many words as possible from the just-presented list in any order. Following presentation of the last item, subjects performed 30 s of arithmetic problems of the form $A + B + C = ?$, where A , B and C were positive, single-digit integers. Subjects were encouraged to solve each problem as quickly and as accurately as possible. In addition to the base rate of \$15, subjects were given a bonus of up to \$15 based on the number of correctly solved problems across distraction intervals. Because the amount of time devoted to the math problems was fixed, a subject may have been presented with, but not responded to, a problem at the end of the distraction interval. Following this distractor task, a tone lasting for 1 s indicated the start of the 90 s recall period. Subjects gave vocal responses that were digitally recorded and processed offline, and were encouraged to attempt recall throughout the duration of the period.

Results

Figure 5 shows the serial position curves for each list type. In addition to the mixed lists and pure once-presented lists simulated with CMR, in this new experiment we included an additional manipulation to control for the possibility that, in mixed lists, subjects may “borrow” time from the second presentation of massed items to rehearse spaced items, leading to elevated recall of the latter. If so, then any observed spacing effect could merely reflect a *rehearsal borrowing* strategy rather than an underlying property of the memory system (for a review see Delaney et al., 2010). In our experiment we included pure massed lists and pure spaced lists, a common method to consider spacing effects unadulterated by rehearsal borrowing (Delaney & Verkoijen, 2009; Hall, 1992; Kahana & Howard, 2005; Toppino & Schneider, 1999). CMR has no mechanism for rehearsal borrowing strategies

and thus this concern only arises in our experimental data. The design of the mixed lists and pure once-presented lists in the experiment were identical to those for CMR simulations.

The arithmetic distractor task attenuated the recency effect, as determined by regressing the recall probability on serial position for the 4 terminal list items. The distribution of the recency slopes was not reliably greater than 0 for any of the list types, Bonferroni corrected $p > 0.2$. To further rule out contributions of serial position effects, we used the same control analyses as described in the CMR results section. First we generated recall sequences based on the distributions of recall probabilities, where recalls were random and independent across items. Because there were many fewer experimental than simulated trials, we generated 100 times as many trials as the observed data. We then calculated the repetition and spacing effects from the randomized recall sequences. In the statistical analyses reported below we subtracted the randomized values from the observed values. Because these differences are usually less than 0, in Figure 6 we present the differences with the mean random list value added to each observed value.

Consistent with previous research, the probability of recall was higher for massed items than once-presented items ($t(34) = 6.23, p < 0.001$), demonstrating a repetition effect. Figure 6 shows recall probability as a function of spacing for items in mixed lists. Subjects also exhibited a spacing effect, reflected as a lower probability of recall for massed items in comparison to repeated items with lag > 0 ($t(34) = 3.80, p < 0.001$).

To evaluate the reliability of the lag effect in the experimental data, we computed the correlation between lag and recall probability for each subject, controlling for serial position effects. In both the pure and mixed lists, the distribution of correlations across subjects was reliably greater than zero (pure: $r = 0.25, p < 0.001$; mixed: $r = 0.21, p < 0.001$). Although this effect has been reported previously (e.g., Delaney & Verkoijen, 2009; Glenberg, 1979; Greene, 1989; Kahana & Howard, 2005; Madigan, 1969; Melton, 1970; Toppino & Schneider, 1999; Verkoijen & Delaney, 2008), it is not always observed (Delaney & Verkoijen, 2009; Hall, 1992; Toppino & Gracen, 1985; Underwood, 1969; Verkoijen & Delaney, 2008; Waugh, 1970).

We next examined CMR's predictions concerning OR scores of once-presented items. Analogous to our control for serial position effects of repeated items, we determined OR scores from randomized recall sequences with recall probabilities matched to the serial position curve of controls lists. We subtracted the OR scores based on these randomized recall sequences from the observed OR scores, and for illustrative purposes added the mean randomized OR score at each lag. Consistent with previous work (Lohnas et al., 2011) and the predictions of the CMR model, we found a significant correlation between OR score and lag in lists of once-presented items, termed the OR score effect ($r = 0.73, p < 0.001$; Figure 7). As predicted by CMR (Figure 3), we also found a significant OR score effect for once-presented items in mixed lists ($r = 0.52, p < 0.001$). In addition, we found a significant OR score effect in massed lists when we treated the presentation of each massed item as a single presentation ($r = 0.56, p < 0.001$). Although the difference between lag = 0 and lag = 1 contributes substantially to these correlations (Figure 7), the OR score effect was still reliable after excluding lag = 0 from the analysis (once-presented lists: $r = 0.53, p < 0.001$;

mixed lists: $r = 0.41, p < 0.001$; massed lists: $r = 0.38, p < 0.001$). These analyses indicate that contextual variability can contribute to recall performance in both pure and mixed repetition lists.

In CMR, contextual variability contributes to both OR score and contiguity effects. The degree to which a subject exhibits a contiguity effect can be quantified by a temporal clustering score (Polyn et al., 2009; Sederberg, Miller, Howard, & Kahana, 2010). A clustering score is calculated by determining the absolute value of the lags between the serial positions of the just-recalled word and the set of not-yet-recalled words at each output position. The observed lag is then ranked amongst all possible absolute lags, and this ranking is normalized on a scale from 0 to 1. Specifically, of all possible rankings the highest ranking is set to 0 and the lowest ranking is set to 1, and the other ranking values are set on a linear scale between the two values. In general, a subject who exhibits stronger temporal organization will make more recall transitions with lower lags and thus higher rankings, and resultingly produce a higher temporal clustering score. Lohnas et al. (2011) found that subjects who exhibit stronger contiguity effects generally exhibit larger OR score effects, thus implicating contextual variability as a factor that influences both measures. Here we consider the relationship between these two measures.¹

For each list type, the temporal clustering scores are reliably greater than those expected by chance yet are less than perfect (all Bonferroni corrected $p < 0.001$). Retrieved context models were originally developed to explain contiguity effects in lists of once-presented items. The present finding of contiguity effects in mixed lists, along with previous findings of contiguity effects in pure spaced lists (Kahana & Howard, 2005), raises the question of whether CMR predicts this contiguity effect in lists with repeated items. Using the same parameters and simulation lists as described above, we found that CMR predicts contiguity effects in mixed lists as well as in pure once-presented lists, at levels comparable to the data (CMR: 0.76, 0.64, data: 0.70, 0.72, respectively; see also Polyn et al., 2009).

Next, we quantified the relationship between the OR score effect and the contiguity effect by calculating the across-subject correlation (combined across once-presented and massed lists; Lohnas et al., 2011). This reliable correlation ($r = 0.38, p < 0.05$) suggests that subjects who encode items according to the assumption of contextual variability (and thus in accordance with CMR) exhibit stronger temporal organization during recall.

Because the study-phase retrieval mechanism applies only to repeated items, CMR assumes that only contextual variability influences the OR score effect of once-presented items, and that both study-phase retrieval and contextual variability contribute to the spacing effect. Thus, just as we used the temporal clustering score to discern the contribution of contextual

¹For lists containing repeated spaced items, the lags for such an item is ambiguous. For instance, the transition between an item presented in serial position 5 to an item presented in serial positions 3 and 9 could be considered a lag of -2 or $+4$. For any transition with an ambiguous lag (i.e. a transition including a repeated item), we randomly selected the lag value from the set of possible lags. We calculated the temporal clustering score in this way for 5 replications of the data, and defined the mean across replications as the clustering score. Another complicating factor in this analysis is that there may be multiple potential recalls at a given minimum lag and thus a particular lag may be included twice in a single transition. To obtain a fair measure of the effect of temporal clustering, as well as maintain a score for repeated items consistent with lists of once-presented items, at each transition a lag could only be listed once in the rankings of possible lags.

variability to OR scores, which consider two items presented once, we can similarly consider the correlation between temporal clustering scores and recall probability of repeated items, which consider one item presented twice. If contextual variability was the only underlying factor driving spacing effects, then we would also expect to see a positive correlation between the lag effect and temporal clustering scores. However, we did not find a significant correlation between these two measures (combined across mixed and pure spaced lists; $r = 0.02$, $p > 0.8$). This nonsignificant correlation would be difficult to explain by assuming that subjects use different temporal clustering strategies across list types, as there is a strong correlation between the clustering score collapsed across control and pure massed lists and the clustering score collapsed across mixed and pure spaced lists ($r = 0.71$, $p < 0.001$).

Given that subjects exhibit significant temporal clustering in lists containing repeated spaced items, yet the degree to which they exhibit temporal clustering does not correlate with the spacing effect, a factor other than contextual variability must also influence the spacing effect. Thus, the lack of a correlation is consistent with the hypothesis that study-phase retrieval interacts with contextual variability to give rise to the spacing effect, as implemented in CMR (see also Benjamin & Tullis, 2010; Glenberg, 1979; Verkoijen et al., 2004; Young & Bellezza, 1982).

Finally, we assessed CMR's prediction concerning transitions between items that follow the same repeated item. When an item initially presented in serial position i is also presented at serial position j , according to CMR the context associated with the presentation at i will be retrieved for the presentation at j (Figure 1C). Because this retrieved context is incorporated into the current context representation, the context from the presentation at i becomes partially associated with $j + 1$ and to a lesser extent $j + 2$. This contextual retrieval mechanism thus implies that items $j + 1$ and $j + 2$ should be associated with a similar context to that of items $i + 1$ and $i + 2$. As such, transitions between items in positions $S_j = \{j + 1, j + 2\}$ and $S_i = \{i + 1, i + 2\}$ should be significantly elevated. For spaced items (lag = 4) in mixed lists, we determined the proportion of times, given that a subject made a transition between an item from an item in S_i or S_j , that they would then transition to an item in S_j or S_i , respectively. To control for the proportion of transitions expected in the absence of repeated items, we assigned the same serial positions used in the mixed lists to corresponding items in the control lists. For each subject, we assigned these positions to 100 random shuffles of the control lists, and took the mean across lists to get a baseline expectation of remote transitions. We found that 34 out of 35 subjects make reliably more of these transitions in mixed lists than expected based on their matched transitions from relabeled control lists (Figure 8), and across all subjects this effect is significant ($t = 8.8$, $p < 0.001$).

Discussion

The context maintenance and retrieval model (CMR; Polyn et al., 2009) embodies two theories of repetition effects: contextual variability and study-phase retrieval. CMR's assumption that the context of an item's first presentation is retrieved during its second presentation (study-phase retrieval) causes an item repeated in a list to be associated with

two, rather than one, effective retrieval cues, thus increasing its chance of being recalled. Based on the assumption that context changes with each presented item (contextual variability), the distinctiveness of these retrieval cues increases with the number of items intervening between a repeated item's two presentations. Thus, CMR predicts that items repeated with larger spacings are more likely to be recalled. At very large lags, the contexts of a repeated item may be sufficiently different, such that increasing the lag even further may not improve recall. Indeed, at large lags increasing the spacing is sometimes detrimental (e.g. Verkoeijen, Rikers, & Schmidt, 2005). Here we considered CMR predictions of the reliable increase in recall probability for shorter spacings.

CMR makes the novel prediction that subjects are more likely to transition between items following a shared repeated item, even if such items are widely separated in the list (Figure 4; Figure 8). This suggests that the second presentation of an item influences encoding of subsequently presented items. CMR also predicts that recalling at least one of two presented items increases with lag (Figure 3; Figure 7). The degree to which subjects showed this effect correlated with their temporal clustering scores (Lohnas et al., 2011) suggesting that both measures can be explained with CMR's assumption of contextual variability. However, the degree to which subjects showed a spacing effect did not correlate with their temporal clustering scores. Because subjects exhibited similar levels of temporal clustering between list types, the lack of a correlation cannot be attributed to a differential contribution of contextual variability. Rather, a second independent process, unique to repeated items, must contribute to this effect. According to CMR, this second process reflects the retrieval of the context associated with the repeated item's first presentation.

It is worth noting that Ross and Landauer (1978) failed to find an OR score effect, and for decades this was taken as evidence against contextual variability theory. In a meta-analysis of six studies, Lohnas et al. (2011) found that OR scores do increase with lag, and suggested reasons for this discrepancy. The current experiment eliminates a possible difference between the two studies: Whereas Ross and Landauer (1978) considered performance of once-presented items in mixed lists, Lohnas et al. (2011) only considered recall of items in pure once-presented or pure massed lists. Thus, one could argue that spaced repetitions have sufficiently clouded the context variability signal such that context was not strictly drifting over time. In CMR terms, the repetition of an item retrieves its prior context, thereby binding items across larger spacings to a similar context (e.g. as in Figure 1C). Thus, Ross and Landauer's (1978) failure to observe an OR score effect could have arisen because they used mixed lists. However, our finding of a significant OR score effect in mixed lists – which contained a higher ratio of repeated to nonrepeated items than the original Ross and Landauer (1978) study – suggests that the presence of spaced items does not necessarily eliminate the OR score effect.

Raaijmakers (2003) presented a version of the Search of Associative Memory (SAM) model that assumed contextual variability theory but did not produce an OR score effect. Because SAM assumes that, irrespective of their spacing, two once-presented items i and j share approximately the same number of contextual features with the retrieval cue, the probability of recalling i given recall of item j is not predicted to increase with their lag. However, Raaijmakers (2003) only defined SAM rigorously for the cued recall paradigm, and thus its

OR score prediction is difficult to compare to CMR's. Given SAM's success in accounting for spacing effects in cued recall, it may be that its instantiation of contextual variability applies to that task but not to free recall. Future work remains to fully characterize the nature of the contextual variability mechanism as well as its role in different types of recall tasks.

In addition to Ross and Landauer's (1978) failure to find an OR score effect, other evidence has accrued against the contextual variability explanation of spacing effects. As reviewed in Delaney et al. (2010), the most common experiment to test contextual variability theory considers memory for items repeated with the same versus different contexts. Contextual variability theory predicts that an item repeated in different contexts should be better remembered than an item repeated in the same context. However, tests of this prediction have yielded mixed results: sometimes spaced items repeated in different contexts are not better remembered than items repeated with the same context (Glanzer & Duarte, 1971; Hintzman, Summers, & Block, 1975; Maskarinec & Thompson, 1976; Postman & Knecht, 1983) and sometimes memory is worse for spaced items repeated in multiple contexts (Postman & Knecht, 1983; Verkoeijen et al., 2004). In contrast, massed items sometimes benefit from being repeated with varied contexts (Glanzer & Duarte, 1971; Maskarinec & Thompson, 1976; Verkoeijen et al., 2004). According to CMR, for spaced items the contextual variability advantage of two different contexts is counteracted by the study-phase retrieval advantage of retrieving the item's prior context. Different types of contextual information could induce different levels of study-phase retrieval, producing either no change or worsened memory for items repeated with the same versus different contexts. For massed items, the similarity in temporal contexts is not as strong an influence as the differences in the experimentally manipulated context. Thus, contextual variability for massed items would be more likely to produce a recall advantage.

Our findings of differences between OR scores (two items presented once) versus recall probabilities of repeated items (one item presented twice) is consistent with CMR's assumption that contextual variability alone cannot account for the spacing effect. For small lags in mixed lists, OR scores exceed recall probability of repeated items, yet OR scores reach a maximum value at lower lags than is typically seen for repeated items. In our implementation of CMR, two unique items have two distinct temporal contexts yet a repeated item is associated with one context—an amalgam of the two presentations. As a result, a repeated item's context does not benefit from improved recall as much as if it were associated with two entirely disparate cues.

Our transition analysis considered how repeating an item served to repeat its context, thus influencing the organization of surrounding items in memory. Howard, Venkatadass, Norman, and Kahana (2007) also considered this question, and found that, in comparison to a list of purely once-presented items, subjects are more likely to transition from a repeated item to an item following either of its presentations. Our results go beyond those of Howard et al. (2007) by demonstrating that the repeated item need not be recalled to influence the organization of its neighbors.

In the experiment reported here, we also found spacing and lag effects in pure lists. We included the pure lists to ensure that the significant spacing effects in mixed lists did not

merely reflect subjects' potential use of a rehearsal borrowing strategy, in which subjects use the time during the second presentation of a massed item to rehearse other spaced items (Delaney et al., 2010). Although Hall (1992) failed to find a spacing advantage in pure lists (which control for rehearsal borrowing), Toppino and Schneider's (1999) replication with a larger range of lags did find an advantage of spaced over massed items, and with slight variations Kahana and Howard (2005) replicated their results. The finding of a spacing effect in pure lists suggests that rehearsal borrowing cannot entirely account for the spacing effect in mixed lists. Here we show how retrieved context theory, as embodied in CMR, could account for the effects of repeated items and once-presented items in mixed lists.

Although we found that retrieved context theory can account for the set of repetition effects reported here, other mechanisms may contribute as well. For instance, another popular theory of repetition effects, deficient-processing theory, assumes that a repeated item's memory trace is retrieved with strength inversely proportional to its lag (Greene, 1989; Hintzman, 1974; Toppino, Fearnow-Kenney, Kiepert, & Teremula, 2009). In the framework of retrieved context theory, this could mean that the amount of context retrieved for a repeated item decreases as a function of its lag (e.g., Pavlik & Anderson, 2005). Future work remains to assess these subtle distinctions.

Our results constrain existing theories of repetition effects by implicating a strong role for a slowly drifting context representation that is updated with each newly presented, repeated, or recalled item. Further, CMR serves to unify memory theories applied to lists of once-presented items and those applied to repetition effects. Recall dynamics of once-presented items and repeated items are traditionally studied separately, or at best, as an interaction between two types of items with different properties. Yet our findings suggest that recall dynamics of both types of items are well-characterized by considering their associated temporal contexts.

Acknowledgments

This research was funded by National Institutes of Health grant MH55687.

References

- Benjamin AS, Tullis J. What makes distributed practice effective? *Cognitive Psychology*. 2010; 61:228–247. [PubMed: 20580350]
- Bower, GH. Stimulus-sampling theory of encoding variability. In: Melton, AW.; Martin, E., editors. *Coding processes in human memory*. John Wiley and Sons; New York: 1972. p. 85-121.
- Cepeda NJ, Pashler H, Vul E, Wixted JT, Rohrer D. Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*. 2006; 132(3):354–380. [PubMed: 16719566]
- Delaney PF, Verkoefen PPJL. Rehearsal strategies can enlarge or diminish the spacing effect: Pure versus mixed lists and encoding strategy. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 2009; 35(5):1148–1161.
- Delaney PF, Verkoefen PPJL, Spigel AS. Spacing and testing effects: A deeply critical, lengthy and at times discursive review of the literature. *Psychology of Learning and Motivation*. 2010; 53(63-147)
- Donovan JJ, Radosevich DJ. A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology*. 1999; 84(5):795–805.

- Estes WK. Statistical theory of spontaneous recovery and regression. *Psychological Review*. 1955; 62:145–154. [PubMed: 14371893]
- Gershman SJ, Moore CD, Todd MT, Norman KA, Sederberg PB. The successor representation and temporal context. *Neural Computation*. 2012; 24(6):1553–1568. [PubMed: 22364500]
- Glanzer M, Duarte A. Repetition between and within languages in free recall. *Journal of Verbal Learning and Verbal Behavior*. 1971; 10:625–630.
- Glenberg AM. Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition*. 1979; 7:95–112. [PubMed: 459836]
- Greene RL. Spacing effects in memory: Evidence for a two-process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1989; 15:371–377.
- Hall JW. Unmixing effects of spacing on free recall. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 1992; 18:608–614.
- Hintzman, DL. Theoretical implications of the spacing effect. In: Solso, RL., editor. *Theories in cognitive psychology: The Loyola symposium*. Erlbaum; Potomac, MD: 1974. p. 77-97.
- Hintzman DL, Summers JJ, Block RA. Spacing judgments as an index of study-phase retrieval. *Journal of Experimental Psychology: Human Learning and Memory*. 1975; 104(1):31–40.
- Howard MW. Scaling behavior in the temporal context model. *Journal of Mathematical Psychology*. 2004; 48:230–238.
- Howard MW, Kahana MJ. Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1999; 25:923–941.
- Howard MW, Kahana MJ. A distributed representation of temporal context. *Journal of Mathematical Psychology*. 2002; 46:269–299.
- Howard MW, Kahana MJ, Wingfield A. Aging and contextual binding: Modeling recency and lag-recency effects with the temporal context model. *Psychonomic Bulletin & Review*. 2006; 13(3): 439–445. [PubMed: 17048728]
- Howard MW, Venkatadass V, Norman KA, Kahana MJ. Associative processes in immediate recency. *Memory & Cognition*. 2007; 35(7):1700–1711. [PubMed: 18062547]
- Kahana MJ. Associative retrieval processes in free recall. *Memory & Cognition*. 1996; 24:103–109. [PubMed: 8822162]
- Kahana MJ, Howard MW. Spacing and lag effects in free recall of pure lists. *Psychonomic Bulletin & Review*. 2005; 12:159–164. [PubMed: 15948289]
- Landauer TK, Dumais ST. A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*. 1997; 104:211–240.
- Loftus GR, Masson MEJ. Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*. 1994; 1:476–490. [PubMed: 24203555]
- Lohnas LJ, Polyn SM, Kahana MJ. Contextual variability in free recall. *Journal of Memory and Language*. 2011; 64(3):249–255. [PubMed: 21379369]
- Madigan SA. Intraserial repetition and coding processes in free recall. *Journal of Verbal Learning and Verbal Behavior*. 1969; 8:828–835.
- Maskarinec AS, Thompson CP. The within-list distributed practice effect: Tests of the varied context and varied encoding hypotheses. *Memory & Cognition*. 1976; 4(6):741–746. [PubMed: 21287005]
- Melton AW. The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Memory*. 1970; 9:596–606.
- Murdock BB. The serial position effect of free recall. *Journal of Experimental Psychology*. 1962; 64:482–488.
- Nelson DL, McEvoy CL, Schreiber TA. The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments and Computers*. 2004; 36(3):402–407.
- Pavlik PI, Anderson JR. Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive Science*. 2005; 29:559–586. [PubMed: 21702785]
- Polyn SM, Norman KA, Kahana MJ. A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review*. 2009; 116(1):129–156. [PubMed: 19159151]

- Postman L, Knecht K. Encoding variability and retention. *Journal of Verbal Learning and Verbal Behavior*. 1983; 22:133–152.
- Raaijmakers JGW. Spacing and repetition effects in human memory: application of the sam model. *Cognitive Science*. May; 2003 27(3):431–452.
- Ross B, Landauer T. Memory for at least one of two items: Test and failure of several theories of spacing effects. *Journal of Verbal Learning and Verbal Behavior*. 1978; 17(6):669–680.
- Sederberg PB, Gershman SJ, Polyn SM, Norman KA. Human memory reconsolidation can be explained using the temporal context model. *Psychonomic Bulletin & Review*. 2011; 18(3):455–468. [PubMed: 21512839]
- Sederberg PB, Howard MW, Kahana MJ. A context-based theory of recency and contiguity in free recall. *Psychological Review*. 2008; 115(4):893–912. [PubMed: 18954208]
- Sederberg PB, Miller JF, Howard WH, Kahana MJ. The temporal contiguity effect predicts episodic memory performance. *Memory & Cognition*. 2010; 38(6):689–699. [PubMed: 20852233]
- Shankar KH, Howard MW. A scale-invariant internal representation of time. *Neural Computation*. 2012; 24:134–193. [PubMed: 21919782]
- Steyvers, M.; Shiffrin, RM.; Nelson, DL. Word association spaces for predicting semantic similarity effects in episodic memory. In: Healy, AF., editor. *Cognitive psychology and its applications: Festschrift in honor of Lyle Bourne, Walter Kintsch, and Thomas Landauer*. American Psychological Association; Washington, DC: 2004.
- Thios SJ, D'Agostino PR. Effects of repetition as a function of study-phase retrieval. *Journal of Verbal Learning and Verbal Behavior*. 1976; 15:529–536.
- Toppino TC, Fearnow-Kenney MD, Kiepert MH, Teremula AC. The spacing effect in intentional and incidental free recall by children and adults: Limits on the automaticity hypothesis. *Memory & Cognition*. 2009; 37(3):316–325. [PubMed: 19246346]
- Toppino TC, Gracen TF. The lag effect and differential organization theory: Nine failures to replicate. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1985; 11:185–191.
- Toppino TC, Schneider MA. Observation: The mix-up regarding mixed and unmixed lists in spacing-effect research. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1999; 25(24):1071–1076.
- Underwood BJ. Some correlates of item repetition in free-recall learning. *Journal of Verbal Learning and Verbal Behavior*. 1969; 8:83–94.
- Usher M, McClelland JL. The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*. 2001; 108(3):550–592. [PubMed: 11488378]
- Verkoeijen PPJL, Delaney PF. Rote rehearsal and spacing effects in the free recall of pure and mixed lists. *Journal of Memory and Language*. Jan; 2008 58(1):35–47.
- Verkoeijen PPJL, Rikers RMJP, Schmidt HG. Detrimental influence of contextual change on spacing effects in free recall. *Journal of Experimental Psychology Learning, Memory, and Cognition*. Jul; 2004 30(4):796–800.
- Verkoeijen PPJL, Rikers RMJP, Schmidt HG. Limitations to the spacing effect. *Experimental Psychology*. 2005; 52(4):257–263. [PubMed: 16302534]
- Waugh NC. On the effective duration of a repeated word. *Journal of Verbal Learning and Verbal Behavior*. 1970; 9:587–595.
- Young DR, Bellezza FS. Encoding variability, memory organization, and the repetition effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1982; 8(6):545–559.

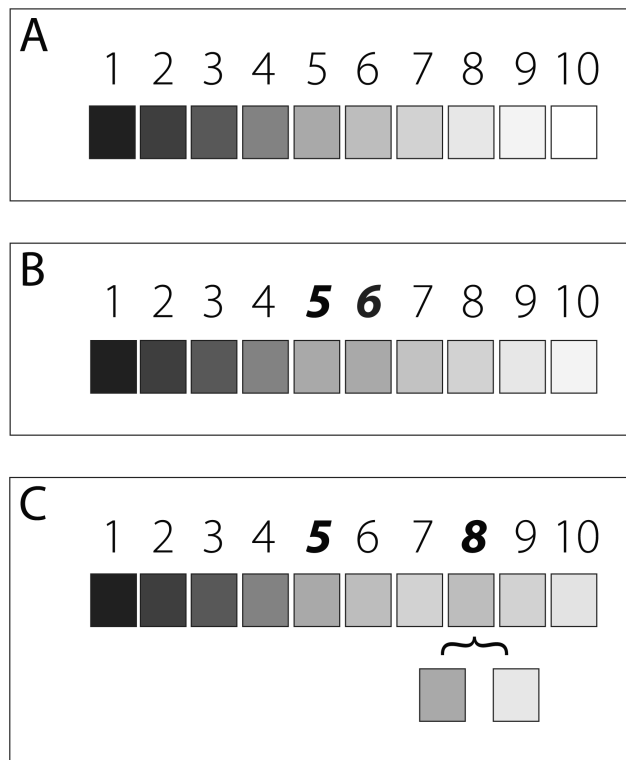


Figure 1. Context representations associated with lists of once-presented and twice-presented items according to retrieved context theory

Each number represents the serial position of a presented item, and below the number is a one-dimensional (grayscale) depiction of context (similar context states appear as similar shades of gray). Serial positions in bold italic indicate that the same item is presented at both positions. **A. List of once-presented items.** Context changes slowly with each presented item. **B. List on once-presented items with a single massed repetition.** The item repeated at position 6 is associated with the retrieved context from its first presentation at position 5. This item is more likely to be recalled because it is associated with 2 retrieval cues. **C. List of once-presented items with a single spaced repetition.** The context of the item repeated at position 8 is an amalgam of the retrieved context from position 5 (left context representation under the bracket) and the slowly drifting context state (right context). The context at position 9 is thus more similar to the context at position 5 than in the case of a list of once-presented items, e.g. Panel A.

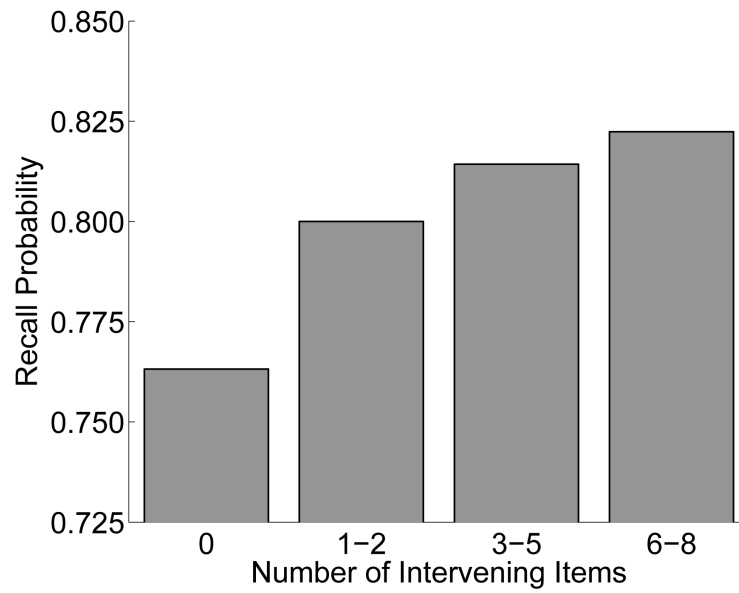


Figure 2. The context maintenance and retrieval model (CMR) predicts the spacing and lag effects in mixed lists

CMR predictions for lists containing a mixture of once-presented items, massed items (repeated at lag = 0) and spaced items (here, divided by short, medium, and long lags).

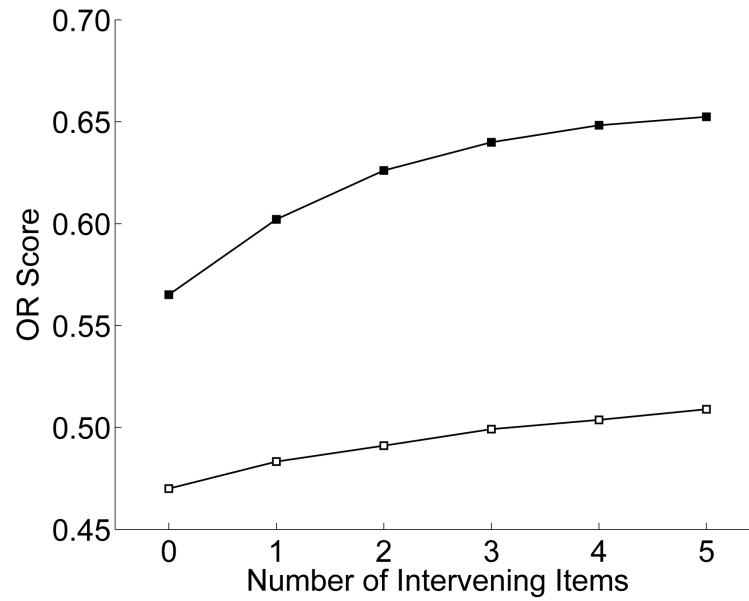


Figure 3. The context maintenance and retrieval model (CMR) predicts the OR score effect CMR predicts that the probability of recalling one item or the other (OR score) increases with the number of items intervening between their presentations. Open squares: OR scores in for once-presented items in control lists. Filled squares: OR scores for once-presented items in mixed lists.

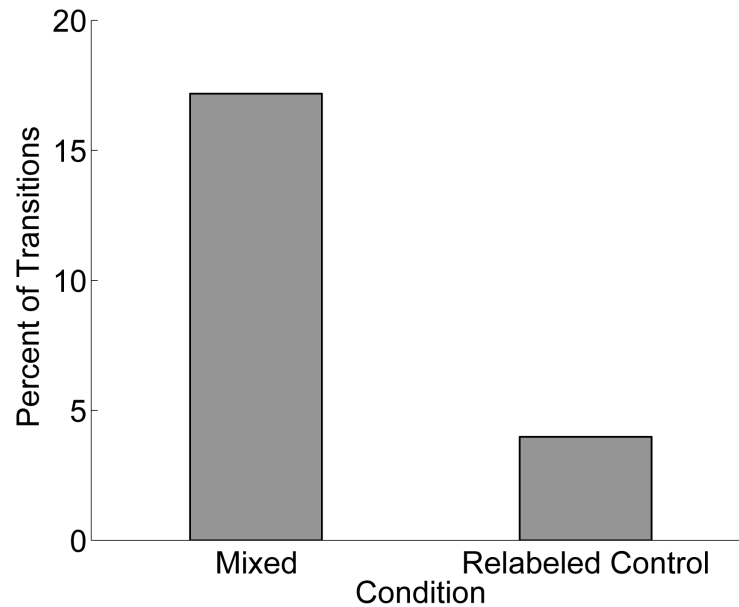


Figure 4. The context maintenance and retrieval model (CMR) predicts that subjects are more likely to make transitions between items that follow a shared repeated item

Left bar: The percent of transitions made between items that follow the presentation of the same repeated item. That is, if an item was presented at positions i and j , and if one of the items was recalled from either $S_j = \{j + 1, j + 2\}$ or $S_i = \{i + 1, i + 2\}$, we then counted the percentage of times that a transition was then made to one of the items in the complementary sets S_i or S_j , respectively. Right bar: The percent of transitions expected by chance. For 100 reshuffled data sets, repeated items in control lists were matched to the same serial positions as the repeated items in mixed lists and the percentage of transitions were counted.

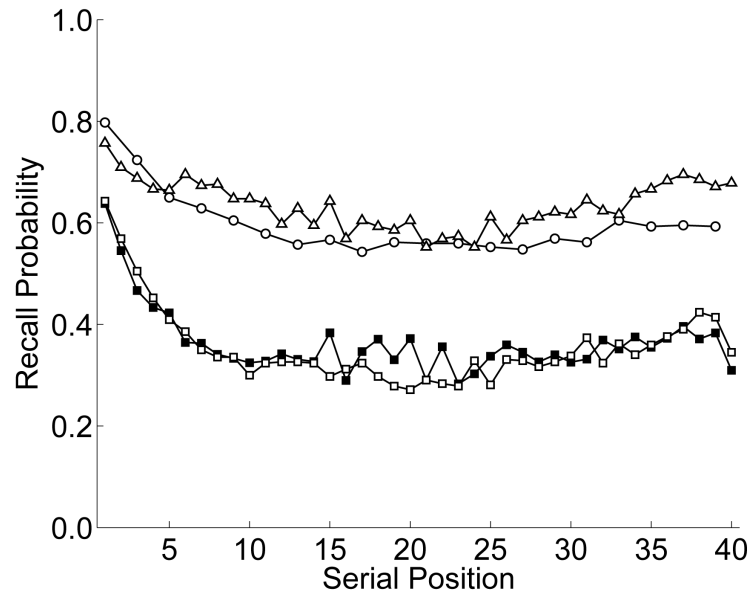


Figure 5. Recall probability as a function of serial position and list type

Open squares: Control lists of once-presented items. Filled squares: Once-presented items in mixed lists. Open circles: Pure massed lists (recall probability for an item in an even serial position n is the same as the recall probability for an item in the corresponding odd serial position $n - 1$, and so for clarity only the odd serial positions from this list type are shown). Open triangles: Pure spaced lists.

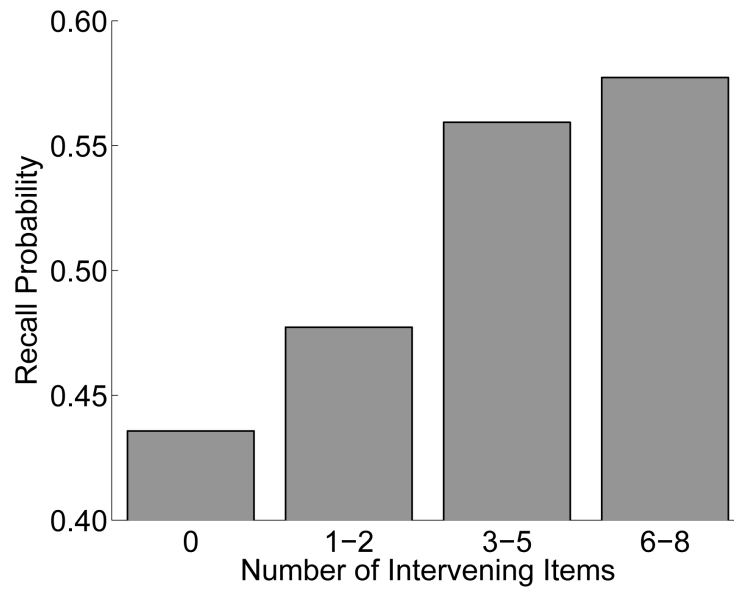


Figure 6. Replication of spacing and lag effects

Recall probability of repeated items in lists containing a mixture of once-presented items, massed items (repeated at lag = 0) and spaced items (repeated with lags 1-8). There is a significant recall advantage for recall of spaced items over massed items (the spacing effect), and recall probability correlates with lag for spaced items.

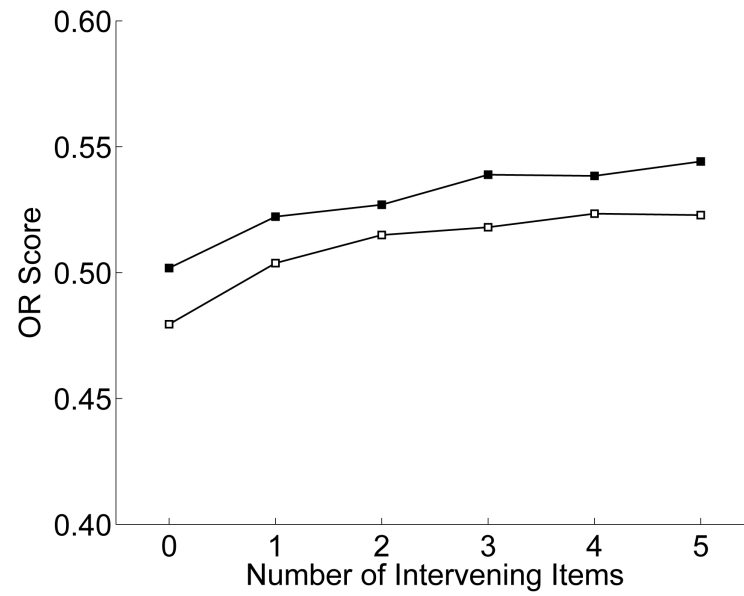


Figure 7. The OR score effect in control and mixed lists

The probability of recalling one item or the other (OR score) increases with the number of items intervening between their presentations. Open squares: control lists of once-presented items. Filled squares: Once-presented items in mixed lists.

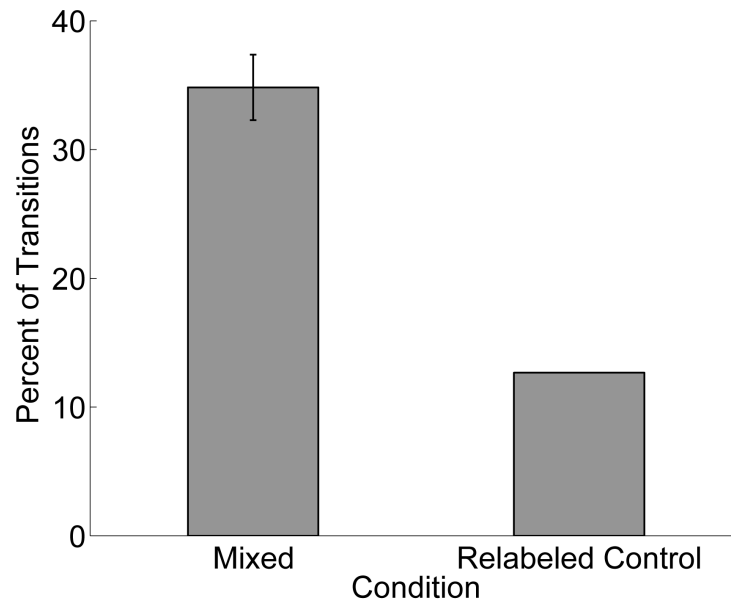


Figure 8. Subjects are more likely to make transitions between items that follow a shared repeated item

Left bar: The percent of transitions made between items that follow the presentation of the same repeated item. That is, if an item was presented at positions i and j , and if one of the items was recalled from either $S_j = \{j + 1, j + 2\}$ or $S_i = \{i + 1, i + 2\}$, we then counted the percentage of times that a transition was then made to one of the items in the complementary sets S_i or S_j , respectively. Error bar represents ± 1 standard error of the mean. Right bar: The percent of transitions expected by chance. For 100 reshuffled data sets, repeated items in control lists were matched to the same serial positions as the repeated items in mixed lists and the percentage of transitions were counted.

Table 1

Parameter values used for the simulations of the Context Maintenance and Retrieval (CMR). Except for β_{dist} , these are the best-fit parameter values of CMR reported in Polyn et al. (2009), originally determined for free recall data from Murdock (1962). This immediate free recall study did not require a value of β_{dist} , which we set from delayed free recall simulations of a predecessor of CMR, the Temporal Context Model with Accumulators (Sederberg et al., 2008).

CMR Parameters	
Parameter	Value
β_{enc}	0.745
β_{rec}	0.357
γ^{FC}	0.581
s	1.80
κ	0.091
λ	0.375
η	0.182
τ	0.242
φ_s	5.39
φ_d	1.41
β_{dist}	0.976