

Expanding the Scope of Memory Search: Modeling Intralist and Interlist Effects in Free Recall

Lynn J. Lohnas
University of Pennsylvania

Sean M. Polyn
Vanderbilt University

Michael J. Kahana
University of Pennsylvania

The human memory system is remarkable in its capacity to focus its search on items learned in a given context. This capacity can be so precise that many leading models of human memory assume that only those items learned in the context of a recently studied list compete for recall. We sought to extend the explanatory scope of these models to include not only intralist phenomena, such as primacy and recency effects, but also interlist phenomena such as proactive and retroactive interference. Building on retrieved temporal context models of memory search (e.g., Polyn, Norman, & Kahana, 2009), we present a substantially revised theory in which memory accumulates across multiple experimental lists, and temporal context is used both to focus retrieval on a target list, and to censor retrieved information when its match to the current context indicates that it was learned in a nontarget list. We show how the resulting model can simultaneously account for a wide range of intralist and interlist phenomena, including the pattern of prior-list intrusions observed in free recall, build-up of and release from proactive interference, and the ability to selectively target retrieval of items on specific prior lists (Jang & Huber, 2008; Shiffrin, 1970). In a new experiment, we verify that subjects' error monitoring processes are consistent with those predicted by the model.

Keywords: memory, free recall, context, list before last, model

Despite the vast stores of memories we accumulate over a lifetime of experience, the human memory system is often able to target just the right information, seemingly effortlessly. Like a powerful “search engine,” our memory system is highly cue dependent, needing the right search terms to find the target information among a sea of structurally similar but contextually distinct items or events. The goal of the present article is to develop a new theory of memory, built upon existing context-based models, which can account for these features of human memory search as observed in the free-recall task. Whereas previous models have frequently made the simplifying assumption that memory search is automatically restricted to a target list, our model explicitly simulates the accumulation of memories across many lists and pro-

vides mechanisms for the selective retrieval of those memories encoded in a given temporal context. This allows us to test our model against experimental findings of proactive interference (PI) and retroactive interference (RI) between lists. Our model also considers the role of semantic memory both in guiding episodic memory retrieval and as a source of interference, depending on the experimental circumstances.

The amazing capacity to retrieve contextually appropriate information during memory search can be seen in many everyday settings, such as when we try to recall the items on our to-do list, or the people to whom we owe a dinner invitation, or the produce we need to pick up at the market. In each case, the memory search engine must be provided with some search terms that help to restrict retrieval to a particular set of memories. For instance, the question “What did you eat for dinner last night?” restricts memories to be associated with a particular context (dinner) but also to a particular semantic category (food).

The set of features surrounding but not comprising the memory itself, termed context, has long roots in the history of memory search. Contextual features may include external factors such as the physical environment and timing of the event, as well as internal factors such as one's inner thoughts. Although context has been a pillar for classic theories of forgetting, spontaneous recovery, and spacing effects (Estes, 1955b; McGeoch, 1932; Underwood, 1945), the role of context in memory gained popularity when Tulving (1972) coined the term “episodic memory” to refer to a memory associated with the spatiotemporal context in which it occurred. A defining feature of episodic memory is that the

Lynn J. Lohnas, Department of Psychology, University of Pennsylvania; Sean M. Polyn, Department of Psychology, Vanderbilt University; Michael J. Kahana, Department of Psychology, University of Pennsylvania.

This research was funded by National Science Foundation Grant 1058886 to Michael J. Kahana and National Science Foundation Grant 1157432 to Sean M. Polyn. We thank David Huber and Yoonhee Jang for sharing their data for Simulation 3. We are grateful to Jonathan Miller and Patrick Crutchley for assistance with designing and programming Experiments 1 and 2, and we thank Kylie Hower, Joel Kuhn, and Elizabeth Crutchley for help with data collection. Model simulation code and simulation data are available at <http://memory.psych.upenn.edu/Publications>.

Correspondence concerning this article should be addressed to Lynn J. Lohnas, Department of Psychology, New York University, 6 Washington Place, 8th Floor, New York, NY 10025. E-mail: ll95@nyu.edu

memory can be retrieved given its context as a cue, and in a complementary way the memory can serve as a cue to retrieve its context. Around the same time, Bower's (1967) multicomponent theory of the memory trace offered a precise definition to the concept of temporal context, linking this idea to the evolution of conditioned elements in Estes' classic stimulus sampling theories (Estes, 1955a). According to this model, contextual representations are composed of many features which fluctuate from moment to moment, slowly drifting through a multidimensional feature space. Whereas previous investigators had noted the importance of temporal coding (e.g., Yntema & Trask, 1963), Bower's theory placed the ideas of temporal coding and internally generated context on a sound theoretical footing.

The Bower model provided the basis for more recent computational models of temporal context and its central role in episodic memory (Howard & Kahana, 2002; Mensink & Raaijmakers, 1988; Murdock, 1997). One such model, Howard and Kahana's (2002) temporal context model (TCM), has been shown to account for a wide range of memory phenomena obtained in the free-recall paradigm (Howard, 2004; Howard & Kahana, 1999, 2002; Howard, Kahana, & Wingfield, 2006; Howard, Venkatadass, Norman, & Kahana, 2007; Sederberg, Gershman, Polyn, & Norman, 2011; Sederberg, Howard, & Kahana, 2008). TCM characterizes the processes involved in the storage and retrieval of memories in terms of associations between item and context representations, with context gradually evolving in response to the information retrieved by the memory system. Polyn, Norman, and Kahana (2009) generalized the class of retrieved temporal context models to accommodate other types of information in the contextual system. Whereas the original TCM assumed that the context representation retrieved by an item is uniquely related to that item until it forms new temporal associations to other items in the course of the experiment, the context-maintenance and retrieval (CMR) model of Polyn et al. (2009) assumes that the context representation retrieved by each item also reflects long-standing semantic as well as source associations.

Like most other models, CMR assumes that only current list ("correct") items exist in memory. Therefore, retrieval in these models can be compared with a search engine that searches only within the most relevant file. Models with this simplifying assumption sidestep the major challenge of how to focus memory search within a broad set of relevant and irrelevant information. Indeed, previous work leaves open the question of whether contextual drift and retrieval processes are sufficient to explain the fundamental problem of list specific memory search (Usher, Davelaar, Haarmann, & Goshen-Gottstein, 2008). Here we present a memory model in which the associations formed between items and context accumulate across all lists in an experimental session. In this way memories accumulate both within and across lists, enabling the model to account for recall of both current list items ("correct recalls") and prior list items ("intrusions"). We show how the present model, termed CMR2, can help to characterize both the mechanisms by which memory search can selectively target a given list, and the ways in which it is affected by interference from items learned in other contexts. We first present an overview of CMR2, then a series of simulations demonstrating that CMR2 can address the key across-list interference effects in free recall. In these simulations, we also test several novel predictions of CMR2, all of which are upheld in the experimental data.

Overview of the Model

Allowing memories to accumulate across many lists poses a major computational challenge to any model of memory search. Specifically, the model must be able to select a small set of target items among a very large set of competitors. People are surprisingly good at this, but they do make predictable errors (e.g., Zaromb et al., 2006). A good model should mimic this basic pattern of human behavior. Moreover, people appear to be able to censor errors when they do come to mind, an idea that served as the basis for the class of generate-recognize models that were once popular (Atkinson & Juola, 1974; Bahrick, 1970; Kintsch, 1970). Despite their achievements, current retrieved context models have no basis for making prior-list intrusions (PLIs), or demonstrating PI effects across multiple lists. They also have no mechanism for targeting specific lists, or censoring responses when their associated contexts are poorly matched to the target context.

Here we introduce a retrieved context model in which memories continually accumulate across lists. Temporal context is used to target retrieval of items learned on a particular list. CMR2 is so named for the continuity between the basic assumptions of contextual maintenance and retrieval with our earlier work (Polyn et al., 2009), as well as to express our view that the innovations introduced here will be necessary for the future development of retrieved context models. A complete description of the model is given in Appendix A, and a summary of model parameters is provided in Table 1.

To illustrate the conceptual principles of CMR2, we provide a simplified example with two lists of five items each. Figure 1 shows the basic structure of the model: Items are represented by the distribution of activations across elements (nodes) of a feature vector, \mathbf{f} , and context is represented by the distribution of activa-

Table 1
Summary of Free Parameters in CMR2

Category	Parameter	Description
TCM base	β_{enc}	Rate of contextual drift at encoding
	β_{rec}	Rate of contextual drift at retrieval
	γ_{FC}	Relative weight of pre-exp. to exp. context (item-to-context)
	γ_{CF}	Relative weight of pre-exp. to exp. context (context-to-item)
TCM-A additions	ϕ_s	Primacy scale factor
	ϕ_d	Primacy decay rate
	κ	Strength of recurrent inhibition
	λ	Strength of lateral inhibition
	η	SD of accumulator noise
CMR addition	s	Semantic scale factor
CMR2 additions	β_{post}^{recall}	Rate of contextual drift between recall and study
	ω	Threshold scale factor
	α	Threshold decay rate
	c_{thresh}	Upper bound for context similarity threshold
Simulation 3 additions	β_{post}^{pause}	Rate of contextual drift between pause and study
	c_{thresh}^{target}	Lower bound for context similarity threshold

Note. Parameters are classified according to whether they were first introduced in the temporal context model (TCM), TCM with accumulators (TCM-A), the context maintenance and retrieval model (CMR), or for the current model, the continuous memory version of CMR (CMR2).

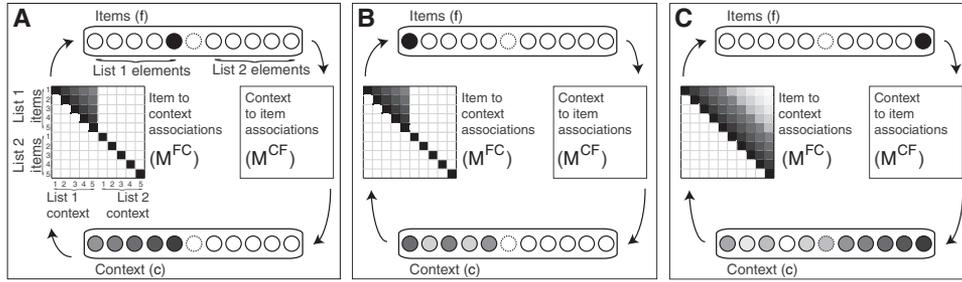


Figure 1. Schematic diagram of the continuous-memory version of the context maintenance and retrieval (CMR2) model during list presentation and recall. Each rounded rectangle represents a vector, and each circle within the rectangle represents an element of the vector. The color of the elements, from black to white, correspond to values of 1 to 0, respectively. Elements corresponding to list items are outlined in solid lines, and outlined in dotted lines are the elements corresponding to the item presented between Lists 1 and 2. (A) Presentation of Item 1 in List 1. (B) Recall of Item 1 in List 1. (C) Presentation of Item 5 in List 2. See text for further details.

tions across elements of a context vector \mathbf{c} . As in previous work, we make the simplifying localist assumption that each item is represented by a standard basis vector (i.e., a vector of unit length with a single nonzero element). These two vectors influence each other via associative matrices M^{FC} and M^{CF} , where M^{FC} stores the strengths of associations from items to contexts and M^{CF} stores associations from contexts to items. Although the orthogonality assumption described above causes all items to be featurally distinct from one another, items can still be related to one another through pre-existing semantic associations in these associative matrices (described below).

Figure 1A shows the state of \mathbf{f} and \mathbf{c} after Item 5 in List 1 was presented. The associated item element is set to 1 in \mathbf{f} , represented by the shading of that element. In addition, this item creates an input to context, \mathbf{c}^{IN} . Context is then updated via the item-to-context association matrix, M^{FC} . For an item i , its associated context state \mathbf{c}_i is calculated as:

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

where ρ_i is a constant ensuring that $\|\mathbf{c}_i\| = 1$ (see Appendix A for the formal definition of ρ_i). According to Equation 1, context is a recency-weighted sum of presented items. The amount by which element values decay with each presented item is governed by the model parameter β_{enc} . A large value of β_{enc} causes context states to decay more quickly. Because List 2 has not yet been presented, elements representing List 2 items are set to zero.

Figure 1A also illustrates the Hebbian learning rule whereby the just-presented item is associated with the previous state of context. In this way, CMR2 associates an item with the temporal context in which it occurs, forming a new episodic memory. These new associations, as well as the context states associated with each of the previously presented items, are represented in M^{FC} .

Once all items in a list have been presented, recall begins by using the current state of context as a retrieval cue. Using context-to-feature associations, each item is assigned an activation based on the sum of its activations across all context states. These activations are used as the starting points for a retrieval competition in which all items race to cross a threshold, following the dynamics of the leaky accumulator model of Usher and McClelland (2001). Although the item with the

strongest initial activation has the best chance of winning, the competition is noisy such that items with lower activations may win.

Whereas previous models assumed that episodic memory was a “tabula rasa,” erased anew before each list presentation, CMR2 assumes that memories accumulate continuously across lists. As such, the decision process governing memory retrieval is not artificially limited only to the items on the current target list. Instead, previous list items compete alongside current list items for retrieval. To allow the most active nonlist competitors to influence retrieval without having to simulate the diffusion for competitors with near zero activation values, we limit the competition to the top l ($= 4 \times \text{list-length}$) competitors.

Because earlier retrieved context models restricted competition to target list items, there was no need to filter out intrusions during recall. We include a “recognize” stage in the retrieval process that filters out inappropriate responses that may be sampled during recall (Anderson & Bower, 1972; Bahrick, 1970; Jacoby & Hollingshead, 1990; Kintsch, 1970; Raaijmakers & Shiffrin, 1980). This idea is consistent with the observation in free recall experiments that people often report thinking of items that they do not overtly recall (Keppel, 1968; Wixted & Rohrer, 1994). In CMR2 terms, suppose an item has just been retrieved and the current state of context is \mathbf{c}_r . The retrieved item creates an input to context, \mathbf{c}_r^{IN} , and updates context as in Equation 1. The input to context from this item is then compared with the current state of context, \mathbf{c}_{r+1} , yielding a similarity value (Anderson & Bower, 1972; Dennis & Humphreys, 2001):

$$u = \mathbf{c}_{r+1}^{\text{IN}} \cdot \mathbf{c}_r. \quad (2)$$

How the value of u is used to filter recalls depends on the experimental procedure. In standard free recall, the item must be from the most recently presented list, and thus its context is expected to match closely with the current context. Thus, if the match is high (i.e., u exceeds a threshold parameter, c_{thresh}), the item is recalled; otherwise, it is filtered out. For other variants of free recall, the modifications to this mechanism are described in more detail in their respective sections.

Regardless of whether an item meets the recognition criterion required for overt recall (see Equation 2), the retrieval of an item

from the decision competition necessarily retrieves its associated context. This retrieved context updates the current context state using the same equation as during list presentation (Equation 1), although the rate of context updating can differ between the encoding and recall periods (β_{enc} and β_{rec} , respectively). This reflects the hypothesis that the rate of context integration may be different depending upon whether a stimulus was externally presented or internally retrieved. In addition, a higher value of β_{rec} impacts recall dynamics in a different way than a higher value of β_{enc} . Specifically, a higher value of β_{rec} causes an item's studied context to be reinstated more completely, pushing out context states of previously retrieved (or presented) items. In contrast, a higher value of β_{enc} leads to weaker representations of previously presented items in the current context state. Thus, irrespective of β_{rec} , a higher value of β_{enc} means that reinstatement of an item's context during the recall period will carry less of the history of presented items.

The updated context representation is then used as the retrieval cue for the next retrieval competition. The recall period is modeled as a series of retrieval competitions that terminates after a fixed number of time steps that reflects the length of the recall period in the experimental studies being simulated.

Whereas previous retrieved context models (e.g., TCM/CMR) made the simplifying assumption that each item could only be recalled once during a given recall period, subjects can and occasionally do repeat items during recall. Subjects may also think of such repetitions, but censor them before recalling them overtly. CMR2 does not exclude the possibility of repetitions during recall. Rather, it permits any item to be retrieved more than once during a recall period. To limit the frequency of such repetitions, previous theorists have suggested that recalled items are temporarily suppressed, making them less likely to compete with nonrecalled items (Burgess & Hitch, 1999; Duncan & Lewandowsky, 2005; Farrell & Lewandowsky, 2002; Henson, 1998; Lewandowsky & Farrell, 2008). In CMR2, we implement a response suppression mechanism by assuming that retrieval of an item results in a temporary increase in its retrieval threshold (see Appendix A for the equations that govern this process).

Figure 1B shows the state of the model after recall of the first item in List 1. The associative strength of List 1 items to the current context no longer indicates how recently each item was presented in the list. Rather, each item's strength is also influenced by context states retrieved during the recall period for List 1. Thus, although context always represents the recency-weighted sum of context states, with recall between lists this does not always correspond to the recency-weighted sum of context states from list presentation.

In CMR2, context changes between lists as subjects transition from a recall mode to a study mode. The degree of this interlist context shift is determined by parameter β_{post}^{recall} , as this shift takes place following each recall period. This idea of a context change as subjects transition between retrieval and encoding modes is supported by a number of prior studies (Aslan & Bäuml, 2008; Pastötter & Bäuml, 2007; Pastötter, Bäuml, & Hanslmayr, 2008; Sahakyan & Kelley, 2002). This context shift is simulated by presenting a new item to the model, and allowing the features of this item to update context using Equation 1 with the context drift parameter β_{post}^{recall} , so named because the item is presented post

(after) recall. This interlist context shift makes items studied in previous lists less accessible, as the temporal context associated with prior list items becomes significantly downweighted. In Figure 1, this extra item is represented by an element with dashed lines in the context and item vectors. It is not associated to context and thus it is not represented in the associative matrix. This item also does not enter the recall competition.

Figure 1C shows the state of \mathbf{f} and \mathbf{c} after Item 5 in List 2 is presented. Here the strength in temporal context of each List 2 item is identical to the strength in temporal context in 1A for each List 1 item in the same serial position. This is because each item's context strength is a function of how recently it was presented. In addition, the item-to-context associations among List 2 items are identical to the item-to-context associations among List 1 items in 1A. List 2 items are more weakly associated with List 1 items because the between-list context shift causes the List 1 items to be more weakly associated with the current state of context.

For simplicity, we show the strength of the association being determined by its associative strength to the current state of context. As elaborated in Appendix A, CMR2 also assumes a primacy gradient of attention such that the change in M^{CF} is greatest for early list items, consistent with the notion that early list items benefit from increased encoding efficiency (Serruya, Sederberg, & Kahana, 2014; Tulving & Rosenbaum, 2006). In our example in Figure 1, we also assumed that associative matrices begin the simulated trial with each item being associated with its context element only (and vice versa for context-to-item associations). In the full implementation of the model, the M^{CF} matrix encodes pre-experimental semantic associations among items (e.g., the *dog* item node is most strongly associated with the *dog* context node, but also has a weaker association with the *cat* node, reflecting the fact that the two words are semantically related). The strengths of these semantic associations were determined by Latent Semantic Analysis (LSA; Landauer & Dumais, 1997). LSA allows one to measure the semantic relationship between two words as the cosine of the angle between the words' representations in a multidimensional model of semantic space. These LSA values are incorporated into the context-to-item association matrix based on the hypothesis that similar items appear often in the same temporal contexts during one's lifetime (Rao & Howard, 2008). The relative contribution of semantic versus experimental (episodic) associations to M^{CF} is governed by the model parameter s .

Simulations

We report four sets of simulations to examine how CMR2 can account for the interactions between prior experimental learning and memory for new items. Whereas retrieved-context models have previously been applied to the dynamics of recall for current list items, in Simulation 1 we show how CMR2 can simultaneously account for recall dynamics of both current-list items and PLIs. First, we ensure that CMR2 can account for intralist recency and contiguity as accurately as its single-list predecessors, verifying that our new version of the model preserves such predictions. We then show that CMR2 can account for the rare yet reliable recall of PLIs, as well as the tendency for PLIs to be recalled from more recently presented lists (Murdock, 1961, 1974; Unsworth, 2008; Unsworth & Engle, 2007; Zaromb et al., 2006). We also examine CMR2's novel predictions regarding contiguity between PLIs.

In Simulation 2, we examine the implications of CMR2's generate-recognize mechanism for subjects' ability to judge whether their recalls were from the current list. In particular, we fit our model to new experimental data obtained using the externalized free recall procedure (e.g., Kahana, Dolan, Sauder, & Wingfield, 2005; Unsworth & Brewer, 2010; Unsworth, Brewer, & Spillers, 2010, 2013; Zaromb et al., 2006). In this task, subjects were instructed to say any word that came to mind during the recall period, and to press a key after any intrusions or repetitions. We compare subjects' performance in this task to those who completed an identical experiment of immediate free recall. We fit CMR2 to data from the externalized free recall sessions and then show that with the same parameters the model can account for data from the immediate free recall sessions.

In Simulation 3, we examine how CMR2 can selectively recall items not from the just-presented list, as in standard free recall, but rather the list before the last (Jang & Huber, 2008; Shiffrin, 1970). It has been argued that models such as CMR, which rely solely on an item-based context representation, will also require a list-specific context to selectively target retrieval of list-before-last items (Usher et al., 2008). In fitting data from the list-before-last paradigm, we rely on the dynamics of the evolving context signal to filter retrievals based on their recency of encoding, rather than assuming the existence of an additional, list-specific context.

In Simulation 4, we show that CMR2 can account for the classic buildup of PI resulting from the semantic similarity between items on a target list and items on previously studied lists. We also show that CMR2 predicts a release from semantic PI when presented with a new list of semantically unrelated words. With the same set of parameters, CMR2 also accounts for the minimal semantic PI exhibited in lists of unrelated items.

Simulation Method

For each of the simulations below, we determined a single set of model parameters that provide a good fit for all of the relevant behavioral measures. To determine the behavioral predictions of a given parameter set, CMR2 was presented with the same series of word lists that were presented to subjects. CMR2 generated a set of recall predictions based on each experimental subject session. The average across CMR2's simulated subjects was compared with the average subject performance. This comparison generated a goodness-of-fit statistic, quantified as the sum of squared errors between model and data, weighted by the *SE* of the data (analogous to a χ^2 goodness-of-fit statistic). A genetic algorithm was used to search the parameter space of the model to find the best-fit parameters that minimized the goodness-of-fit statistic (see Appendix B), and these parameters are reported in Table 2. Of the 14 model parameters used to fit data from the standard free-recall task, 10 were inherited from TCM and CMR: the contextual drift rates during encoding and recall (β_{enc} , β_{rec}); the semantic strength between items (s); the strength and decay of the primacy gradient (ϕ_s , ϕ_d); the parameters of the decision competition (κ , λ , and η); and the relative strengths of pre-experimental and experimental associations (γ_{FC} , γ_{CF}). The new parameters in CMR2 are the contextual drift rate between lists (β_{post}^{recall}), the context comparison threshold (c_{thresh}), and the parameters that control the increase in decision threshold for retrieved items (ω, α). Two additional pa-

Table 2
Best-Fit Parameters of CMR2

Category	Parameter	1	2	3	4
TCM base	β_{enc}	0.520	0.481	0.182	0.635
	β_{rec}	0.628	0.329	0.754	0.937
	γ_{FC}	0.425	0.647	0.568	0.500
Primacy	γ_{CF}	0.895	0.780	0.356	0.669
	ϕ_s	1.41	1.30	6.75	1.99
	ϕ_d	0.990	0.393	0.866	0.100
Accumulator	κ	0.313	0.098	0.081	0.190
	λ	0.130	0.143	0.018	0.134
	η	0.393	0.412	0.087	0.353
CMR addition	s	1.29	1.87	2.48	2.47
CMR2 addition	β_{post}^{recall}	0.803	1.00	0.745	0.836
	ω	11.9	7.29	19.8	11.1
	α	0.679	0.794	0.974	0.627
Simulation 3 addition	c_{thresh}	0.074	0.431	0.335	0.185
	β_{post}^{pause}	—	—	0.972	—
	β_{post}^{target}	—	—	0.044	—
	c_{thresh}	—	—	—	—

Note. Numbers correspond to each of the simulation analyses described in detail in the simulations section. 1: Kahana et al. (2002). 2: Externalized free recall experiment. 3: Jang and Huber (2008). 4: Loess (1967). CMR2 refers to the continuous-memory version of the context maintenance and retrieval model. See Appendix B for an explanation of the algorithm used to determine these parameters.

rameters are required to simulate the list-before last paradigm, as described in Simulation 3 below.

Simulation 1: The Effects of Prior Experience on Episodic Recall

When asked to freely recall items from a just-presented list, subjects will occasionally recall items from previously studied lists. These PLIs have been the subject of study because they reflect the PI of prior list learning on current list recall (e.g., Melton & Irwin, 1940; Melton & von Lackum, 1941). An understanding of when and how previous memories interfere with current memories is one of the classic puzzles of verbal learning. In this simulation we show how CMR2 can simultaneously account for major within-list phenomena observed in immediate free recall and for PI effects.

In free recall of “unrelated” word lists, PLIs are relatively rare, a fact that led some early researchers to assume that each item is uniquely tagged to its list (Anderson & Bower, 1972; Bahrick, 1970; Postman, 1976, but see Postman & Hasher, 1972). Consistent with this view, models of recall have frequently assumed that all items presented in the same list were associated to a list-context representation that only changes between lists (e.g., Anderson & Bower, 1972; Farrell, 2012; Lehman & Malmberg, 2009; Sederberg et al., 2011; Sirotin, Kimball, & Kahana, 2005). However, such implementations of list tags address neither how people can access memories of a particular period of time, nor how memories can ideally be associated with both a time scale operating within each list as well as a time scale across lists. Below, we show how CMR2 can account for the relative infrequency of PLIs without requiring list tags. Briefly, the context shift in between lists helps to weaken previous list items in context and thus distinguishes the temporal context of the current list from previous lists. In this way, the post-recall context shift not only helps to differentiate the

current list from previous lists, but also helps to make each list distinct. This between-list disruption, combined with CMR2's assumption that representational strengths in temporal context decrease as a function of list recency, allows CMR2 to predict that PLIs are relatively rare.

In addition, PLIs exhibit a strong recency effect, being much more likely to come from recent than remote prior lists (Murdock, 1961, 1974; Unsworth, 2008; Unsworth & Engle, 2007; Zaromb et al., 2006). This illustrates how retrieval of recent items is supported at time scales much longer than that of a single list. Here we show that the CMR2 mechanism used to predict intralist recency can also be used to generate this PLI recency effect: More recently presented items are represented more strongly in context, and thus are more likely to be recalled. This demonstrates how CMR2 can use context representations to explain recency on multiple time scales.

The strongest test of any model is whether it can make a novel prediction that is borne out in the data. The contextual-retrieval process in CMR2, coupled with the continuity of memory across lists, leads to a striking prediction about the contiguity effects observed for successively recalled PLIs. CMR2 predicts that in the rare cases where subjects commit successive PLIs, they will exhibit strong contiguity effects, both within and across lists. As described below, analyses of a large database of prior free recall studies demonstrate both the within and across list contiguity effect for PLIs predicted by CMR2. Just as the ability to predict intralist recency and PLI recency provides a parsimonious explanation of across-list effects without list tags, the ability of CMR2 to predict intralist and interlist contiguity further validates this model assumption.

Results. We fit the CMR2 model to data from an immediate free recall study in which each subject completed 30 lists of 10 common nouns (Experiment 1 of Kahana, Howard, Zaromb, & Wingfield, 2002). We chose this study for two reasons. First, it recorded the identity of specific PLIs, thereby allowing us to examine the PLI-recency effect. Second, this study did not manipulate any other aspects of the free-recall task, thus making it a good testbed for the basic assumptions of the model.

Current-list recalls. At the beginning of the recall period in immediate free recall, end-of-list context serves as the retrieval cue. Because current-list items have stronger representations in context, their recall is favored over prior-list items.

CMR2 accounts for the within-list recency and within-list contiguity effects in the same way as previous context-based models

(Howard & Kahana, 2002; Polyn et al., 2009; Sederberg et al., 2008). The recency effect refers to subjects' tendency to recall recently studied items first, and to recall those items with a higher probability than midlist items (Figure 2A, B). The contiguity effect refers to subjects' tendency to successively recall items studied in nearby list positions. This can be seen in the probability of recalling item from serial position $i + \text{lag}$ immediately following recall of item i , conditional on the availability of item $i + \text{lag}$ as a valid recall. This conditional-response probability function, known as the lag-CRP, is shown in Figure 2C. The within-list lag-CRP illustrates the asymmetry in the contiguity effect, with forward transitions favored over backward transitions (Kahana, 1996).

Context-based models predict recency because the time-of-test context at the start of the recall period overlaps with the contexts associated with recent list items. Contiguity arises in these models because the context retrieved by an item combines with the current context, which is then used to cue the next recall. The forward asymmetry effect arises because an item's pre-experimental context is incorporated into temporal context only after the item is presented. Consequently, a presented item has contexts more similar to, and thus stronger associations with, items presented after its presentation. In summary, CMR2 can account for recall initiation, recall transitions, and the overall shape of the serial position curve.

Prior-list intrusions. Subjects infrequently, yet reliably, commit PLIs during free recall. In the Kahana et al. (2002) study, subjects recalled, on average, 0.25 PLIs per trial (see Table 3). These PLIs tend to come from more recent lists, reflecting an across-list recency effect. We control for the availability of PLIs from more recent lists by only considering PLIs from lists where PLIs of any list-lag can be recalled (Zaromb et al., 2006). Here we consider PLIs recalled in List 4 and later, as a PLI of list-lag = 3 (the maximum list-lag considered here) cannot be recalled until List 4. With the same parameters used to generate the within-list effects described above, CMR2 also captures these properties of PLIs (see Table 3).

We examined CMR2's novel predictions regarding the successive recall of PLIs. We first examined the across-list lag-CRP for successively recalled PLIs, which is defined as the probability of transitioning from a PLI from list i to a PLI from list j as a function of the list lag $j - i$ (across-list CRP; Howard, Youker, & Venkatasass, 2008; Unsworth, 2008). For instance, if a subject recalled a PLI from List 8 followed by a PLI from List 11, this would represent a list lag of 3. To obtain stable predictions from the model, we simulated each of the sessions from Experiment 1 of

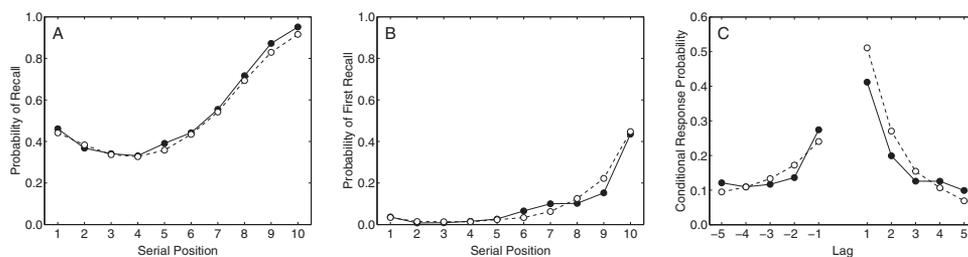


Figure 2. CMR2 predictions of intralist recency and contiguity in immediate free recall. CMR2 predicts intralist effects of free recall (unfilled circles). (A) Serial position curves. (B) Probability of first recall. (C) Conditional response probability as a function of lag. Data from Kahana et al. (2002), Experiment 1 (filled circles).

Table 3
Prior-List Intrusions in Immediate Free Recall

Measure	Data	CMR2
PLI per trial	0.253 (0.035)	0.287
PLI, list-lag = 1	0.411 (0.054)	0.410
PLI, list-lag = 2	0.103 (0.024)	0.072
PLI, list-lag = 3	0.064 (0.018)	0.046

Note. Data are from Experiment 1 of Kahana et al. (2002). SEM are shown in parentheses. PLI = prior-list intrusion; CMR2 = continuous-memory version of the context maintenance and retrieval model. PLIs of a particular list-lag refer to the proportion of PLIS that correspond to that lag. All analyses exclude each subject's first three lists to allow for equal opportunities to make PLIs of each list-lag.

Kahana et al. (2002) 10 times using the parameters reported in Table 2. We compared the model predictions with experimental data aggregated across the studies reported in Table 4. Because participants rarely commit PLIs in succession, testing this theoretically important but subtle prediction necessitated a meta-analysis of data from multiple experiments.

In both the model and experimental data, successive PLIs tended to come from nearby lists, demonstrating an across-list contiguity effect (Figure 3A, B). CMR2 predicts this across-list contiguity effect for the same reason that it predicts within-list contiguity between current-list items: Recall of an item reinstates its associated context, which in turn increases the probability that the next recalled item came from a neighboring list position. We also examined the lag-CRP at the level of items for successive PLI pairs that were originally presented in the same list (Figure 3C, D). CMR2 predicts a within-list contiguity effect for these successive PLIs, again because of its retrieved-context mechanism. Because CMR2 has noise-free representations of items and context, the qualitative prediction of the model overestimates both the within-list and across-list contiguity effects.

CMR2 mechanisms controlling prior-list intrusions. As we have shown above, CMR2 can account for a wide range of data concerning recall of PLIs. By this theory, the same context-based mechanisms that give rise to correct recalls, recency, and contiguity also give rise to these recall errors. We now discuss the major mechanisms in CMR2 that influence recall of PLIs.

Table 4
Studies Included in Conditional Response Probability (CRP) Meta-Analyses

Citation	<i>N</i>	<i>N</i> across	<i>N</i> within	Lists per session	List-length
Bridge (2006)	119	62	23	18	25
Golomb, Peelle, Addis, Kahana, and Wingfield (2008)	36	12	3	36	10
Kahana, Howard, Zaromb, and Wingfield (2002), Exp. 1	30	15	8	30	10
Kahana, Howard, Zaromb, and Wingfield (2002), Exp. 2	25	5	3	20	10
Lohnas, Polyn, and Kahana (2011)	61	38	23	24	24
Polyn, Norman, and Kahana (2009)	45	6	3	17	24
Sederberg et al. (2006)	48	22	19	16	15
Sederberg, Miller, Howard, and Kahana (2010)	27	13	8	16	16

Note. *N* = Number of subjects who participated in the study. *N* across = number of subjects included in the across-list CRP analysis. To be included in this analysis, a subject must have recalled at least one pair of successive PLIs. *N* within = number of subjects included in the PLI-CRP calculated for pairs of successively recalled PLIs from the same lists, i.e., the lag-CRP calculated for those items recall within the same list. To be included in the within-list PLI-CRP analysis, a subject must have recalled at least one pair of successive PLIs in which both PLIs were originally presented in the same list.

1. Representational strengths in temporal context decrease as a function of list recency. PLIs are less likely to be recalled simply because they were not presented in the most recent list. The overlap in temporal context between PLIs' encoding context and the time-of-test context will be less than the overlap between a current list item and the time-of-test context. Even for a PLI retrieved on a previous list, this retrieval took place before the presentation of the current list, and thus such a PLI would not have as strong of a representation in context in comparison with current list items. The further back in time that a PLI was last retrieved (or presented), the weaker feature strength this item will have in the decision competition. Thus, it is less likely to be recalled. This structure leads the CMR2 model to predict the PLI-recency effect.
2. The rate of context drift between lists (β_{post}^{recall}). The context shift in between lists helps to weaken previous list items in context, and thus distinguishes the temporal context of the current list from previous lists. In this way, the post-recall context shift not only helps to differentiate the current list from previous lists, but also helps to make each list distinct. The mechanism implementing between-list contextual change shares certain characteristics with a context disruption mechanism found by Polyn et al. (2009) to be necessary to account for the behavioral effects of within-list shifts in task context.
3. The generate-recognize mechanism. The context associated with a retrieved item is compared with the current state of context (Equation 2). If the retrieved item's associated context does not exceed the similarity threshold (c_{thresh}), then the item will not be recalled. The temporal contexts of PLIs are less similar to the current state of context, and thus are more likely to fail this criterion than current-list items. In general a higher value of c_{thresh} decreases recall of PLIs.
4. The SD of the noise in the decision process (η). A higher value of η increases the likelihood that items with weaker activations will be retrieved. Because

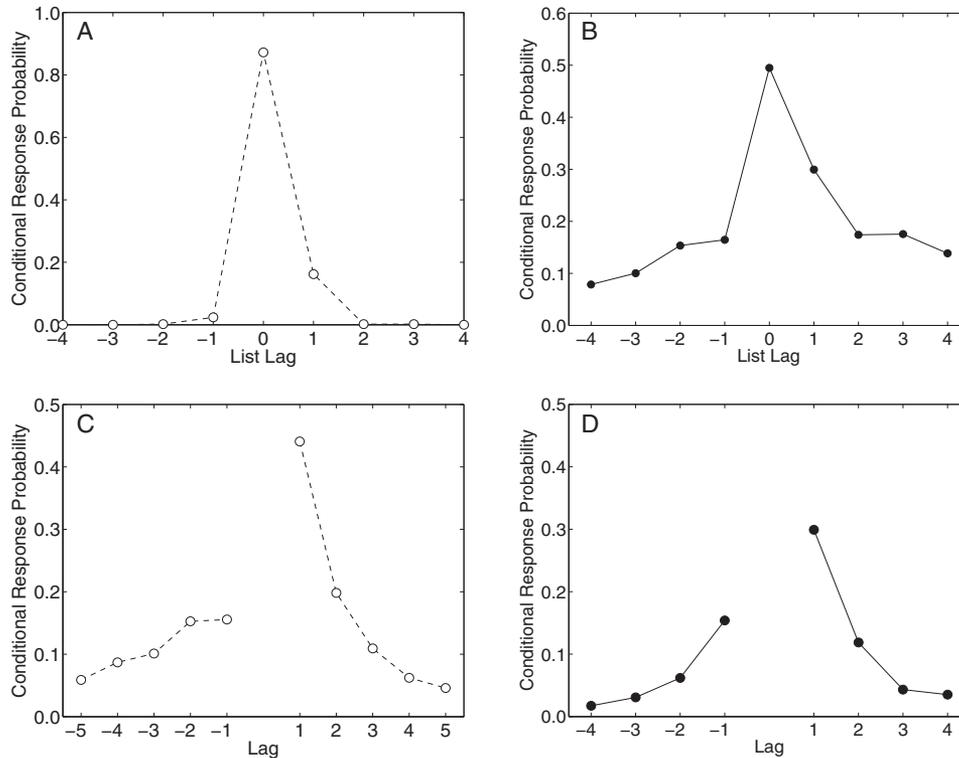


Figure 3. CMR2 predictions and meta-analyses of long-range contiguity in free recall. (A) CMR2 predicts that two PLIs recalled successively are more likely to be presented in nearby lists. (B) A meta-analysis of studies is consistent with CMR2 predictions. (C) Conditional response probability as a function of lag, for two prior-list intrusions (PLIs) recalled successively from the same list. CMR2 predicts that the two PLIs were more likely to have been presented in neighboring serial positions in their original list. (D) Experimental data exhibit similar trends to CMR2's predictions. Data from the meta-analysis were originally reported as Bridge (2006); Golomb, Peelle, Addis, Kahana, and Wingfield (2008); Kahana, Howard, Zaromb, and Wingfield (2002); Lohnas, Polyn, and Kahana (2011); Polyn, Norman, and Kahana (2009); Sederberg, Gauthier, Terushkin, Miller, Barnathan, and Kahana (2006) + replication, Sederberg, Miller, Howard, and Kahana (2010). Where applicable, we only considered experimental manipulations of free recall in younger adults.

PLIs are represented more weakly in the end-of-list context cue, in general a higher value of η increases recall of PLIs.

- The strength of semantic associations between items (s). PLIs with strong semantic associations to recently retrieved items are more strongly activated by the current state of context, and thus are more likely to be recalled. Subjects also exhibit this tendency to transition from a current list item to a strong semantic associate on a prior list (Kahana, 2012; Zaromb et al., 2006). A higher value of s increases the influence of semantic organization in recall output order, and thus increases recall of PLIs.
- Recall of one PLI facilitates recall of other PLIs (Unsworth et al., 2013; Zaromb et al., 2006). Once a PLI is recalled, its associated contexts are reinstated, strengthening the representations of other PLIs presented nearby in time to the just-recalled PLI. These neighboring PLIs will have greater feature strengths in

the subsequent decision competition, and thus will have a greater chance of being recalled. This property gives rise to CMR2's predictions regarding contiguity between successively recalled PLIs (see Figure 3). Just as CMR2 predicts intralist and interlist recency using the same model mechanism, CMR2 also predicts contiguity both within and across lists from its core assumption of using retrieved context to guide recall.

Simulation 2: Using Context for Error Monitoring

Here we use the externalized free recall (EFR) procedure to assess the generate-recognize theory as embodied in CMR2 (Atkinson & Juola, 1974; Bahrick, 1970; Kintsch, 1970; Postman, 1976). In this paradigm, subjects are instructed to say aloud all words that come to mind while performing free recall and to press a key immediately after the recall of an item they believe was not on the most recent list (indicating a "rejection"; Kahana et al., 2005; Unsworth & Brewer, 2010; Unsworth, Brewer, & Spillers, 2010, 2013; Zaromb et al., 2006). By encouraging subjects to use a very low criterion for recalling items that come to mind, the EFR

instruction helps to reveal words that are *generated* but that would not be overtly recalled under standard task instructions. By asking subjects to further reject items that were not on the studied list, the EFR procedure can be used to identify which items were *recognized* as having been studied in the appropriate list context.

Previous EFR studies have yielded a set of results consistent with generate-recognize theory. Subjects recall reliably more PLIs in EFR than in standard free recall, and they further reject the majority of their verbalized PLIs (Kahana et al., 2005; Unsworth & Brewer, 2010; Unsworth et al., 2010, 2013; Zaromb et al., 2006). Although subjects' recall of current-list items is unaffected by the EFR instruction, they do occasionally reject correct items (Kahana et al., 2005; Unsworth & Brewer, 2010; Zaromb et al., 2006).

Because CMR2 uses a context-comparison criterion to determine whether to recall a retrieved item, it is straightforward to simulate the EFR paradigm by simply allowing the model to recall all of the retrieved items and then having the model reject any item that fails the context comparison criterion. Any retrieved item, irrespective of whether it is rejected, updates context. Thus, CMR2 assumes that EFR does not fundamentally change the processes that govern free recall (and context updating).

Below we test predictions of CMR2 that are specific to the nature of its generate-recognize mechanism, namely that items retrieved less recently have less overlap with the current temporal context, and thus have a higher chance of being rejected. This not only leads to a higher rejection probability for PLIs than for correct items, but also that PLIs from more distant lists are more likely to be rejected (Unsworth et al., 2010). We present two new experiments supplemented with simulations to show how CMR2 accounts for these data. First, we present an experiment and simulation of EFR. We then present a CMR2 simulation using the same set of parameters to make predictions concerning a new experiment with identical methods to the EFR experiment except that subjects perform standard immediate free recall (IFR). A complete description of the experimental methods is provided in Appendix C. CMR2's ability to explain both IFR and EFR, with minimal changes to the model across paradigms, suggests that subjects are using the same core memory processes across tasks.

Results. Consistent with previous studies, subjects recalled more PLIs under the EFR instruction, and rejected a majority of these errors (see Table 5).¹ CMR2 predicts this result because items that were retrieved less recently have less overlap with the current temporal context, and thus have a higher chance of being rejected. Following this logic, CMR2 also predicts that PLIs with larger list-lags have a higher probability of rejection. Replicating the finding of Unsworth et al. (2010), we performed a paired *t* test between probability of rejection at list-lag = 1 and larger list lags (in the range 2–5) for the 84 subjects who rejected PLIs at both sets of list-lags. The mean of the distribution for the rejection probability at list-lag = 1 ($M = .80$) was reliably lower than the rejection probability for greater list-lags ($M = .85$), $t(83) = 2.15$, $p < .05$. In other words, items from one list back were more likely to be endorsed as members of the most recent list, compared with items from more distant lists.

A comparison between the experimental data for the IFR and EFR manipulations of the experiment is shown in Figure 4. Similarities between EFR and IFR are indicative of the fact that the EFR procedure makes explicit the implicit memory search processes taking place, rather than forcing a new recall strategy on

subjects. Although subjects make many more overt intrusions in EFR, the recall of correct items did not differ reliably between IFR and EFR sessions, two-sample $t(141) = 0.09$, $p > .5$, suggesting that the increased recall of intrusions in EFR did not hinder the recall of correct items. In addition, the recency advantage in initiation of recall and probability of recall is preserved between the two tasks. Lastly, properties of recalled intrusions are preserved as well: Subjects recall as many PLIs in IFR as the number of nonrejected PLIs in EFR, two-sample $t(141) = 1.42$, $p > .1$.

In both tasks, CMR2 makes qualitatively accurate predictions regarding probability of recall for correct items and PLIs (see Table 5). Although we calculated the best-fit parameters based on subjects' performance in EFR, CMR2 can use the same parameters to predict subjects' performance in IFR because it assumes that the same cognitive processes underlie both tasks. For instance, CMR2 captures the within-list recency effect in IFR and EFR (see Figure 4) for the same reasons as in Simulation 1: At the beginning of the recall period, temporal context is a recency-weighted sum of current list items. In fact, we see consistencies in parameter values between this simulation and Simulation 1, even though here the best-fit parameters were fit using EFR data. Although statistical tests cannot be easily conducted on the differences between parameter values across simulations, in Table 2, each of the critical parameters outlined in Simulation 1 has best-fit values for Simulations 1 and 2 that fall in a similar part of the range of possible values.²

Simulation 3: Retroactive and Proactive Interference in the List-Before-Last Paradigm

In standard free recall, subjects recall items from the most recent list. In the list-before-last paradigm, subjects recall items from the list studied before the most recent list (Jang & Huber, 2008; Lehman & Malmberg, 2009; Sahakyan & Hendricks, 2012; Shiffrin, 1970; Unsworth et al., 2013; Unsworth, Spillers, & Brewer, 2012; Ward & Tan, 2004). That is, following presentation of list n (the intervening list), subjects recall items from list $n - 1$ (the target list). Shiffrin (1970) introduced this paradigm in a landmark article that challenged the classic understanding of RI as being caused by two factors: (a) Extinction, or unlearning, of relevant associations caused by the learning of new associations; and (b) competition between new and old associations, commonly referred to as response competition (Keppel, 1968; Melton & Irwin, 1940; Melton & von Lackum, 1941; Postman & Underwood, 1973). In opposition to predictions of the extinction component of RI, Shiffrin (1970) found that the length of the intervening list had no effect on recall of items in the target list (Jang & Huber, 2008; Sahakyan & Hendricks, 2012; Unsworth, Spillers, & Brewer, 2012; Ward & Tan, 2004). Shiffrin interpreted this finding as evidence for retrieval failure rather than unlearning (see Tulving & Psotka, 1971, for a related result in the case of recall of categorized lists). Although the list-before-last paradigm is quite challenging for

¹ Under the EFR instruction, subjects also showed similar trends for extralist intrusions (recall = 1.4 per list; rejection probability = 0.62) and repetitions (recall = 0.89; rejection probability = 0.61). Because these types of errors are currently beyond the explanatory scope of CMR2, we do not consider them here.

² In the search algorithm, the possible ranges for these parameters were: $\beta_{post}^{recall} \in [0.1, 1]$; $c_{thresh} \in [0.01, 0.5]$; $s \in [0.5, 3]$.

Table 5
Memory Performance in IFR and EFR Experiments

Measure	IFR		EFR	
	Data	CMR2	Data	CMR2
Recall probabilities				
Correct	10.6 (0.38)	10.7	10.6 (0.02)	10.6
PLI	0.08 (0.01)	0.17	0.68 (0.08)	0.61
Rejection probabilities				
Correct	—	—	0.01 (0.001)	0.01
PLI	—	—	0.79 (0.03)	0.76

Note. SEM are shown in parentheses for the experimental data. IFR = immediate free recall; EFR = externalized free recall; PLI = prior-list intrusion; CMR2 = continuous-memory version of the context maintenance and retrieval model. See [Appendix C](#) for experiment methods.

subjects, they nonetheless succeed at recalling an order of magnitude more items from the target list than from the intervening list ([Jang & Huber, 2008](#); [Sahakyan & Hendricks, 2012](#)).

More recent studies using the list-before-list paradigm have shown that the act of recalling list $n - 2$ following list $n - 1$ protects list $n - 1$ from the interference effects caused by list n items. Replacing the recall period between list $n - 1$ and list n by a brief pause, [Ward & Tan \(2004\)](#) found the expected decrease in target list recall with increasing length of the intervening list. In a series of experiments that varied the nature of the task that subjects were given between lists, [Jang and Huber \(2008\)](#) found that between-list tasks requiring episodic memory retrieval replicated Shiffrin's finding that RI did not depend on the length of the intervening list. These results not only support the notion that stronger RI effects reflect a loss in accessibility of relevant cues, but also underscore the role of episodic memory retrieval in influencing RI.

Here we simulate the list-before-last paradigm with CMR2 to characterize how memory retrieval influences the accessibility of a particular list context. We present a simulation study of Experiment 1 of [Jang and Huber \(2008\)](#), in which subjects studied a series of lists varying in target list-length, intervening list-length, and task performed between lists.³ We show that CMR2 is able to account for these experimental manipulations with a single parameter set. The retrieved context mechanism of CMR2 can explain RI in list-before-last paradigm without an explicit list-tagging mechanism.

CMR2 provides a framework to contrast memory search processes in standard free recall with list-before-last recall. Whereas in standard free recall the time-of-test context serves as an effective retrieval cue, recall of the list before the last requires one to search memory for an effective retrieval cue. Explicit simulations of the list-before-last paradigm (e.g., [Jang & Huber, 2008](#); [Lehman & Malmberg, 2009](#)) assume that the target list is associated to a list-specific, directly accessible context (see also [Davelaar, Usher, Haarmann, & Goshen-Gottstein, 2008](#)) and include a parameter determining the probability of reinstating the target-list context. The value of this parameter differs when a pause follows the target list as opposed to when a recall period follows the target list. Although these models can account for the presence or absence of RI, they do not specify the mechanism by which the reinstatement of target-list context occurs. Here we explicitly consider the role of context reinstatement in retrieving and maintaining the target-list context.

The list-before-last paradigm allows for further characterization of between-list context shifts. Whereas this paradigm requires the model to access context states from the prior list, in standard free recall it is detrimental to keep such context states salient in memory. Thus, we might expect the change in context between lists to be weaker in the list-before-last paradigm than in standard free recall. In contrast to a pause between lists, a recall period between lists serves to impose context change between the end of one list presentation and the beginning of the next list presentation. Below, we consider how the best-fit parameters differ with respect to standard free recall, and consider whether the context drift rate must vary based on the between-list task. It has been established that recall between lists serves to distinguish lists from one another ([Pastötter, Schicker, Niedernhuber, & Bäuml, 2011](#); [Szpunar, McDermott, & Roediger, 2008](#); [Tulving & Watkins, 1974](#)), and thus recall between lists may not require as large of a context shift in comparison with a pause. This raises the possibility that the degree of context change can be altered in a task-dependent manner, potentially by an executive control system, an idea we return to in the General Discussion.

The generate-recognize mechanism introduced in [Equation 2](#) is critical for CMR2 to capture the behavioral dynamics of the list-before-last paradigm. This mechanism allows CMR2 to filter out responses from the intervening list until an item associated with the target list context is retrieved. When an item is retrieved, the context retrieved by that item is compared with the active context state. If these two representations are too similar, the generated item is rejected as likely coming from the most recent list. In other words, retrievals exceeding a similarity threshold are omitted (i.e., $\mathbf{c}^{\text{IN}} \cdot \mathbf{c} > c_{\text{thresh}}$). Another threshold, $c_{\text{thresh}}^{\text{target}}$, ensures that retrievals whose contextual similarity falls below this threshold are rejected, as these items are likely from a list preceding the target list. In modeling the list-before-last paradigm, we assume that recall initiates with the first retrieved item that meets the context similarity criteria defined above. Note that this item may not necessarily be a target list item. Nonetheless, CMR2 assumes that this item is from the target list and that retrieval of this item's context serves to cue the next recall. After this first putative target

³ In contrast to the reported 120 subjects included in [Jang and Huber \(2008\)](#), here we only consider the 106 subjects for whom detailed recall dynamics were recorded, allowing for sufficient comparisons with CMR2.

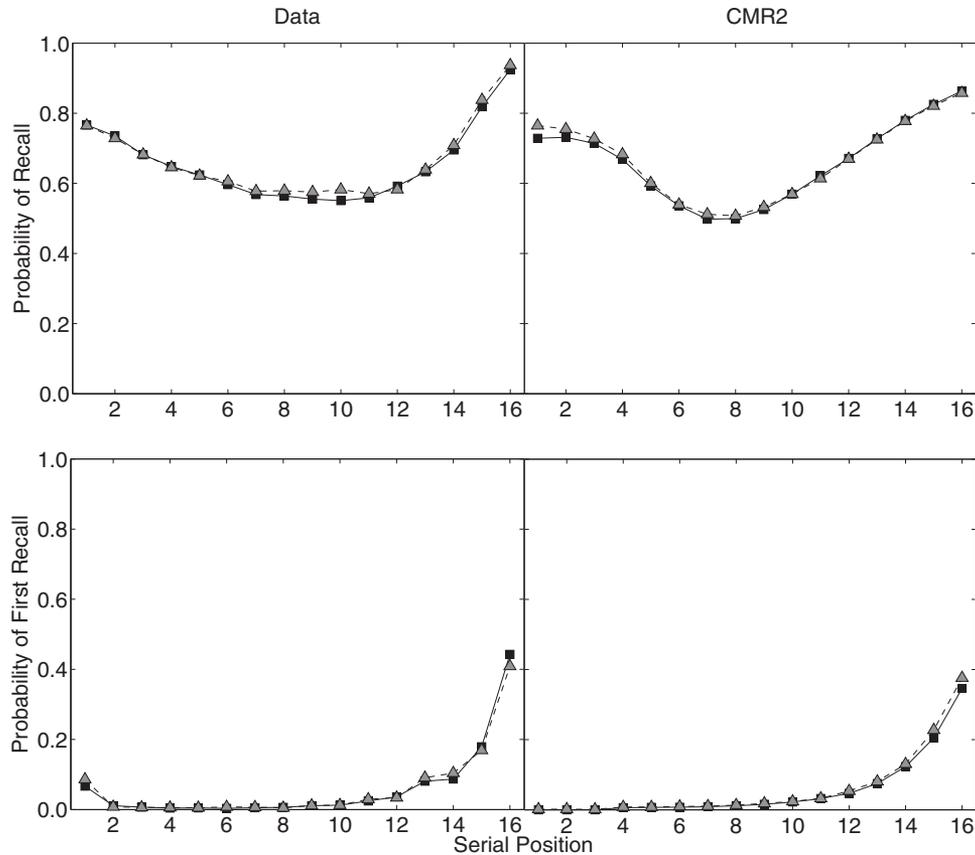


Figure 4. Experimental results (left) and CMR2 simulations (right) of the immediate free recall experiment (gray triangles) and nonrejected items of the externalized free recall experiment (black squares). Across the two experiments (consisting of different sets of subjects), the results between the two experiments are relatively similar. Top row: Serial position curves. Bottom row: Probability of first recall.

list item is recalled, the context comparison mechanism changes to the standard generate-recognize mechanism introduced in our simulations of immediate free recall (retrievals are omitted if they satisfy $e^{IN} \cdot c < c_{thresh}$).

Lastly, in this simulation we test a novel prediction of CMR2 based on its generate-recognize mechanism. Specifically, CMR2 predicts that the proportion of recalled intervening-list intrusions increases with output position. While intervening list items are likely to be rejected in the initial part of memory search, as output position increases, it is more likely that the model has shifted to the context comparison mechanism associated with target list recall. This decision rule is more permissive, as retrieved context must simply exceed c_{thresh} . Now that the model has discovered and reinstated target list context, intervening list items will no longer trigger the extremely high context similarity values that allowed them to be rejected at the beginning of search. In a sense, intervening-list intrusions are analogous to PLIs in free recall, inasmuch as both types of intrusions may meet the context-comparison criterion even though they are from an incorrect list. Furthermore, each retrieval is a potential opportunity for an error, and thus the number of opportunities for errors increases with output position.

Results.

Target-list recalls. Figure 5 shows the experimental data and model predictions regarding the proportion of items recalled from

the target list. As in Experiment 1 of Jang and Huber (2008), we consider the influences of three factors: target list-length (6 or 24 items), intervening list-length (6 or 24 items), and task performed between lists (recall of the target list or a brief pause). In the experimental data, the length of the intervening list produced substantial RI when there was only a pause after the presentation of the target list. CMR2 predicts this effect because each studied intervening list item causes temporal context to drift farther from the target-list context. Thus, with a longer intervening list-length, it is more difficult for CMR2 to reinstate the target-list context. This is consistent with the claim that RI results from decreased accessibility of relevant cues rather than passive unlearning (McGeoch, 1932; Shiffrin, 1970; Tulving & Psotka, 1971).

When subjects were given a recall test between successive lists, increasing the length of the intervening list did not reliably reduce recall of target-list items, $t(105) = 1.80, p > .05$; reanalysis of data from Jang & Huber, 2008. CMR2 accounts for this approximate invariance in target list recall because, during the recall period after the target list, context states of target-list items (list $n - 1$) and earlier-list items ($n - 2$) are retrieved, which serve to reduce the influence of the intervening-list items when the subject is later trying to remember the target-list items. This prediction of CMR2 is not specific to our best-fit parameter set; in a separate search of the parameter space we could not find a set of parameters that led

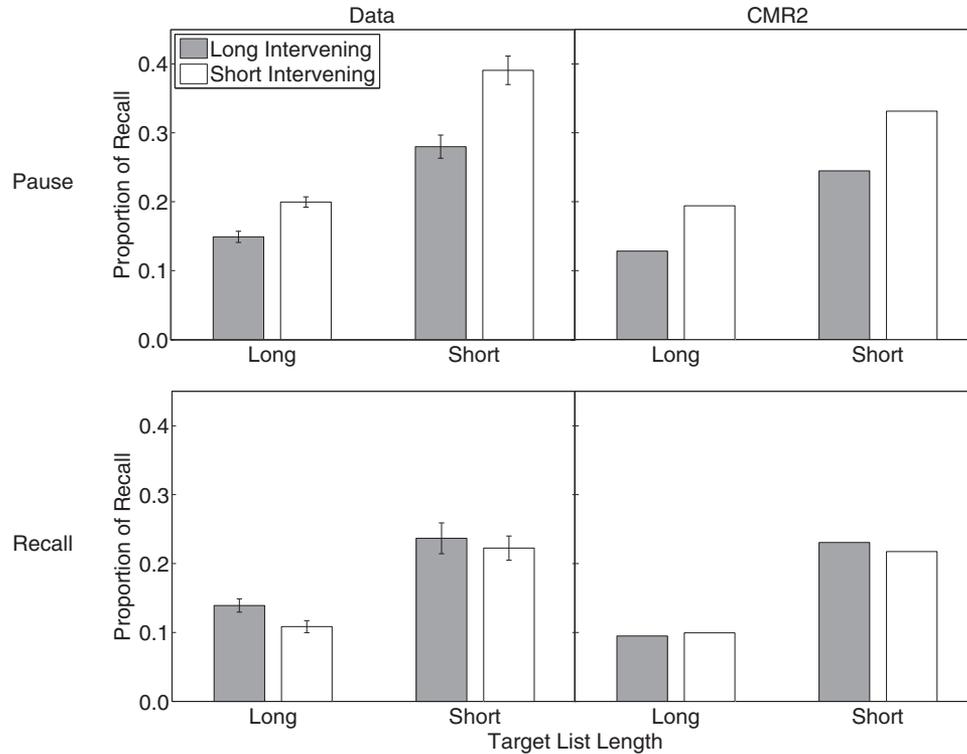


Figure 5. Target-list recalls in the list-before-last paradigm. In the list-before-last paradigm, subjects are asked to recall items from the list before the just-studied list (the target list). The left panels show data from Experiment 1 of Jang and Huber (2008) and the right panels show CMR2 simulations of the experimental data. Both the length of the target list and the length of the list intervening between the target list and its recall period were manipulated. Top panel: With a brief pause between the target list and intervening list, recall of target-list items is negatively affected by increasing intervening list-length. Bottom panel: With recall between lists, target-list recalls are unaffected by intervening list-length. Regardless of the manipulation between lists, proportionally fewer target-list items are recalled with increasing target list-length.

CMR2 to predict an effect of intervening list-length on recall of target list items, assuming that recall levels generally matched those of the experimental data (see Simulation 3a, Appendix B).

To offer some intuition for the model's behavior, Figure 6 shows the similarity between each item's context and the current state of context in each of the two between-list conditions: pause and recall. When there is recall between the target and intervening lists (i.e., recall of list $n - 2$, which we term the earlier list), a subset of items from target and earlier lists are retrieved. Because CMR2 assumes that the contexts retrieved from these items are incorporated into the current context, and the rate of context retrieval in this paradigm is relatively high ($\beta_{rec} = 0.754$), these retrieved items benefit from being more strongly associated to the current context. An item's retrieved context is also strongly correlated with the contexts of its neighbors, and thus items presented nearby on the list to a particular retrieved item may also boast increased association to the current context. Thus, with recall between lists, earlier list and target list items are prominent in context after presentation of the intervening list. Of course, whether there is recall between lists relies not only on the model's attempt to retrieve items, but also its ability to do so. Using the simulation results of the best-fit parameter set, we examined recall of a target list as a function of the number of retrieved items in the previous recall period.

We found that, so long as at least one item was retrieved during that prior recall period, irrespective of what list it came from, CMR2 yielded a null-effect of intervening list length, as seen in the data. This suggests that, in the setting of standard list-before-last paradigm, the retrieval of a single item updates context enough to increase the strength of earlier and target list items in context for the subsequent recall period. We return to this point in the General Discussion.

Irrespective of the task between lists, target list recalls change with target list-length. The list-length effect, whereby the proportion of items recalled from a list decreases with list-length (Murdock, 1962), is prominently observed in Figure 5. The CMR model of Polyn et al. (2009) predicted the list-length effect because the decision competition was comprised solely of current-list items. Increasing the list-length increased the number of items competing for recall, thus increasing the amount of lateral inhibition that each item receives from all other items. However, in CMR2 the same number of items enters the decision competition each time, raising the question of how it accounts for the list-length effect. Whereas the number of competing items is always the same, the support for those items in the competition is not always the same. Significantly, the long target lists have more items associated with a similar contextual cue than the short target lists. Thus, when target-list context is reinstated, target list items have stronger

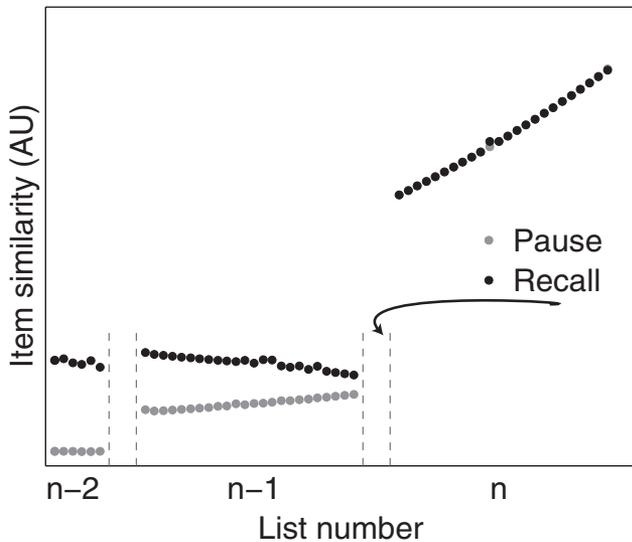


Figure 6. Item similarity and retrieval in the list-before-last paradigm as a function of task between lists. Similarity of earlier list (list $n - 2$), target list ($n - 1$), and intervening list items (n) at the end of presenting the intervening list. With recall between lists, earlier list, and target list items benefit from stronger associations to the current context because of retrieval of those items during the recall period.

activation values in the decision competition than items from other lists, and with longer target list length there will be more such target list items. A decision process including more items with strong activations is functionally equivalent to a decision process with more items competing. As such, proportionally fewer target-list items are recalled with a longer target list.

To better understand the model's ability to account for this complex pattern of behavior, we next examine the parameter values obtained by our optimization procedure. As mentioned previously, we allowed interlist contextual drift to vary across the recall and pause conditions. This was based on the intuition that subjects who must distinguish items from different lists without the benefit of an interlist recall period may use cognitive strategies designed to increase the shift in context during the pause interval (Hintzman & Block, 1971; Sahakyan & Kelley, 2002). Consistent with this idea, we found that the best-fit value of the context shift parameter was higher with a pause between lists ($\beta_{post}^{pause} = 0.97$) than with recall between lists ($\beta_{post}^{recall} = 0.75$). The importance of a lower value for β_{post}^{recall} is also supported by the fact that the value of β_{post}^{recall} is lower in this simulation than the standard free recall simulations ($\beta_{post}^{recall} = 0.803, 1.00, 0.836$ for Simulations 1, 2, and 4, respectively), suggesting that it is more critical for temporal context to not drift too far from the target list, even with the presentation of the intervening list. Whereas the list-before-last paradigm requires subjects to remember items from previously presented lists, thus requiring a lower context drift rate between lists, in free recall there is no such need to remember items from previous lists. To the contrary, in free recall it is more advantageous to weaken the context representations of prior-list items to discourage their retrieval during recall of the just-presented list.⁴

In comparing the estimated model parameters when fitting the list-before-last paradigm to the parameters obtained in standard

free recall of the current list, we found that two additional parameters had best-fit values that diverged substantially across the simulations. The lateral inhibition term in the decision competition ($\lambda = 0.018$) was lower for this simulation ($\lambda = 0.130, 0.143, 0.134$ for Simulations 1, 2, and 4, respectively). This decreased inhibition serves two purposes. First, it makes it more likely that items other than those with the strongest associations to the current temporal context—including target-list items—will be retrieved. Second, it means that intervening-list items are only able to exert RI effects when they are prominent in the recall competition, as in the pause-between-lists condition. The parameter value controlling the primacy effect ($\phi_s = 6.75$) was also much higher for this simulation, as the value of ϕ_s was substantially lower (and rather similar) across the other three simulations ($\phi_s = 1.41, 1.30, 1.99$ for Simulations 1, 2, and 4, respectively). As a result, target list items benefiting from stronger primacy are more prominent in the retrieval competition, allowing the model to more easily transition to the target list.

Intervening-list intrusions. With its generate-recognize mechanism, CMR2 can match subjects' low rate of intervening-list intrusions (see Figure 7). At the beginning of the recall period, this mechanism helps to prevent recall of retrieved intervening-list intrusions by requiring a retrieved item's context to be dissimilar to the current context state; later in recall, after the first putative target-list item is recalled, the mechanism requires a retrieved item's context to be similar to the current (target) context state.

Consistent with the experimental data, CMR2 predicts that the proportion of intervening-list intrusions is greater with recall between lists (see Figure 7). With a pause between lists, CMR2 covertly retrieves more intervening-list intrusions before discovering the target list. These items are assigned a higher threshold for subsequent retrievals (Equation A8), reducing the set of recallable intervening items.

We next consider CMR2's novel prediction that the proportion of recalled intervening-list intrusions increases with output position (irrespective of the task between lists). As CMR2 recalls more items, it becomes increasingly likely that it will mistakenly recall an intervening-list intrusion. More specifically, we considered the number of intervening-list recalls at each output position divided by the total number of intervening and target-list recalls at each output position. CMR2 predicts that the proportion of intervening-list intrusions should increase with output positions, and this prediction is borne out in the experimental data as well. Means of proportion of intervening-list intrusions for output positions 1, 2, 3, and 4 were 0.084, 0.102, 0.121, and 0.135, respectively, and in a paired t test between output positions 1 and 4, the proportion of

⁴ Indeed, when we constrained the model to require the same value of β_{post}^{recall} across conditions, the algorithm search settled on a value substantially higher than when the value could vary across conditions ($\beta_{post}^{recall} = 0.91$), presumably because it was searching for a value that also served to distinguish lists in the pause condition. As a result of this higher value, the representation of the target-list context in the current context was too weak and this variant of CMR2 substantially underpredicted the proportion of target-list recalls in the recall-between-list condition.

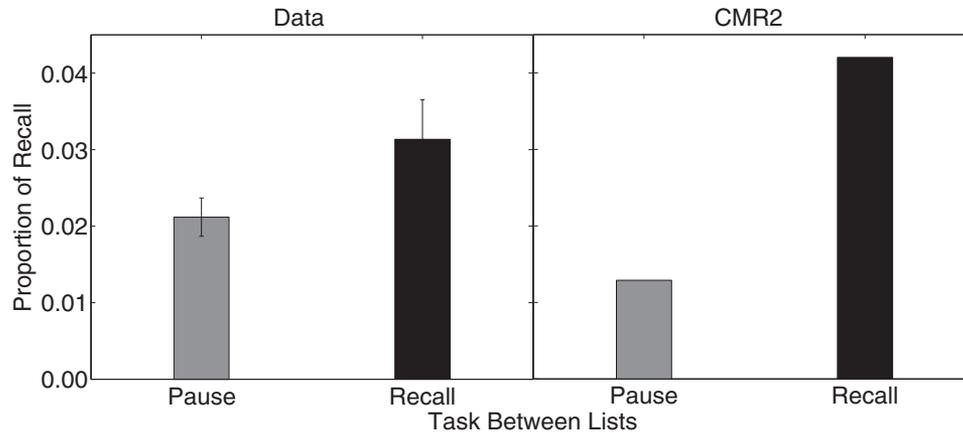


Figure 7. Intervening-list recalls in the list-before-last paradigm. Fewer intervening-list intrusions are recalled when there is no recall between lists in comparison to when there is recall between lists. Nonetheless, recall of intervening-list intrusions is an order of magnitude smaller than recall of target-list items. Data are from Jang and Huber (2008), Experiment 1.

intervening-list intrusions was reliably less for output position 1 than 4, $t(104) = 2.93$, $p < .005$.⁵

Simulation 4: Buildup and Release From Proactive Interference

It is well established that the semantic organization of knowledge, built over a lifetime of experiences, can exert a powerful influence on the way we remember specific events. This is seen in free recall experiments in the way subjects cluster their responses on the basis of semantic similarity. Specifically, after recall of a given list item (e.g., cat) subjects are more likely to transition to recalling a semantically related item (e.g., rabbit) than a dissimilar item (e.g., cello). This positive influence of semantic structure on episodic recall has been accounted for in a number of models of free recall, including those based on retrieved context (e.g., Polyn et al., 2009) and those based on the interaction of short-term and long-term memory mechanisms (e.g., Sirotin et al., 2005).

In the classic analysis of PI in memory, semantic similarity can exert a negative influence on recall performance. This is seen in the buildup of PI as subjects study and attempt to recall lists of semantically related items (e.g., Keppel & Underwood, 1962; Underwood, 1957; Wickens, 1970; Wickens, Born, & Allen, 1963). When subjects are presented with a series of lists that share a common semantic feature (e.g., category), with each subsequent list subjects accumulate more PI from previous lists, which leads to worsened memory for the current list and increased intrusions from prior lists. Although subjects must rely on precise temporal information to recall items only from the most recent list, the interaction between temporal and semantic information in memory search leads to more errors. When a subject is then presented with a list of items unrelated to the previous lists (e.g., new category), the subject exhibits a “release” from PI whereby PLIs are minimal and correct recall is much greater.

Here we examine how CMR2 can account for the major findings of the release-from-PI free recall study presented in Loess (1967). In this experiment, 120 subjects performed delayed free recall on each of 24 three-item lists. Each list contained words from the

same semantic category. After the presentation of the three study items, subjects were given a demanding distractor task in which they had to rapidly recite five six-digit numbers as they appeared visually one at a time. They were then given 10.5 s to recall the three studied items. Subjects were equally divided into the experimental and control groups. In the experimental group, every set of three consecutive lists was from the same category, whereas in the control group successive lists were from different categories (Figure 8A).

In CMR2, as in some earlier versions of the model (e.g., Polyn et al., 2009) semantic similarity can be represented by assuming that semantically related items have been associated with one another’s temporal contexts (Rao & Howard, 2008), and thus are represented in association matrix from context to items. Based on these associations, when CMR2 retrieves a particular item from memory, this supports recall of items not only with similar temporal contexts to the just-retrieved item, but also items with similar semantic contexts. Based solely on this assumption of semantic relations among items, we tested whether CMR2 could account for the semantic and release from PI effects. Thus, we assessed whether the version of CMR2 used in standard free recall (Simulation 1) could account for the buildup and release from PI results without any additional modifications.

Because we do not have the actual lists used in the original study, we could not simulate the subtle variations in semantic similarities among all of the simulated words; rather, we made the simplifying assumption that words from different semantic categories have very low similarities (0.05 on a 0–1 scale) and words from the same category have very high similarities (0.90 on a 0–1 scale). None of the simulation results below depend on these precise values as long as the same category word pairs have much higher similarities than the different category word pairs.

⁵ One subject was excluded for not contributing any recalls at one of the output positions, leading to an undefined value for proportion of intervening-list recalls.

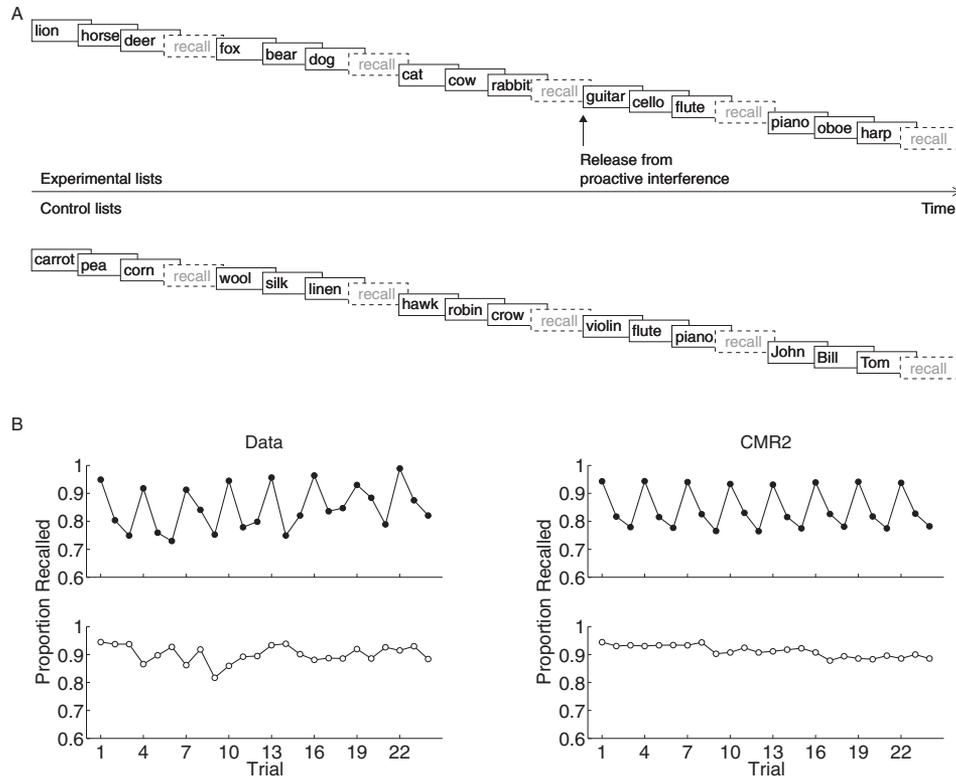


Figure 8. Release from proactive interference (PI) in lists of categorized words. (A) Experiment design. The experimental group was presented with lists of three items, each drawn from the same semantic category (sample words were used in the original study). After three lists from the same semantic category, subjects were then presented with a new set of three lists from a different semantic category. In the control group, lists were presented randomly with respect to category, except that lists drawn from all eight categories were presented before the same category was presented again. (B) Recall performance as a function of trial. Top: Release from PI in experimental lists when a list is presented from a new semantic category. Bottom: No build-up of PI in control lists. Left: Data from Loess (1967). Right: Predictions of the continuous-memory version of the context maintenance and retrieval model (CMR2).

Results. The experimental subjects exhibited a build-up of PI across successive same-category lists, as seen in their declining recall performance (experimental group, Figure 8B, top left panel). CMR2 matches this decline in recall performance for same-category lists as seen in the top right panel of Figure 8B. This is because when an item on a categorized list is retrieved from memory, for the subsequent recall competition this encourages recall of other items that share similar temporal contexts to the retrieved item (from the current list) as well as items that share similar semantic contexts (from prior lists). Thus, CMR2 is more susceptible to interference from PLIs that share semantic (category) information with items from the just-presented list. Because recall is competitive and the recall period is limited, recall of current-list items is attenuated due to interference from these same-category PLIs.

After a category switch, subjects in the experimental group demonstrated a release from PI as seen in their improved performance after every third list. CMR2 also matches this release from PI because if an item is retrieved from such a list, only items from the just-presented list have strongly overlapping temporal and semantic contexts. Thus, recall on these lists is primarily restricted to current-list items.

Recall in the control group exhibited minimal PI effects (Figure 8B, bottom left panel). With the same set of parameter values as for the simulated experimental group data, CMR2 accounts for the relatively constant performance across successive lists in the control group (Figure 8B, bottom right). As in the case of release-from-PI lists, PLIs for the control group share minimal semantic and temporal contextual information and are less liable to produce PI.

In this simulation, none of the critical parameter values were strikingly different from those of Simulations 1 and 2 (see Table 2). This suggests that it is the nature of the semantic relations among items, rather than any specific parameter value, that distinguishes CMR2's performance in standard free recall versus the release-from-PI paradigm.

General Discussion

The emergence of the classic short-term memory (STM) paradigms in the early 1960s, and the concurrent prominence of the serial position effect and other within-list memory phenomena, shaped the development of memory models throughout subsequent decades. The single-list study-test method became the standard

system for studying episodic memory in the laboratory. As a result, many of our leading memory models have been developed, tested, and evaluated primarily on their ability to account for within-list memory effects. This focus on within-list phenomena contrasts with the use of memory in our daily lives, which clearly involves interactions between memories formed in temporally disparate contexts, and the ability to select information learned in a particular context.

Whereas retrieved context models of episodic memory have had considerable success in accounting for the classic within-list memory effects, these models have made the simplifying assumption that memory is reset at the start of each list, and that recall reflects competition among memories learned only within the target list context. Here we present a continuous-memory version of the context maintenance and retrieval model, to extend the explanatory scope of retrieved context models beyond the realm of within-list memory phenomena. We showed how our new model, CMR2, can account for recency and contiguity effects both within and across lists as well as across-list PI and RI effects. CMR2 accounts for these effects using a single slowly changing context representation that serves as the recall cue. In addition, CMR2 makes new, testable predictions regarding recall phenomena and their underlying properties, as we summarize below.

We first examined CMR2's predictions in a standard immediate free recall experiment. CMR2 predicts recency, exhibited both in serial positions of correct items (see Figure 2) as well as in list-lag of PLIs (see Table 3). Earlier versions of CMR predicted within-list recency because the context state used to cue recall is a recency-weighted sum of presented items; CMR2 predicts across-list recency for the same reason. Although in the current manuscript we only consider CMR2 predictions of serial position effects in immediate free recall, CMR2 nests a version of the model that can account for the attenuation of recency in delayed free recall and long-term recency in continual distractor free recall (Sederberg et al., 2008).

As with any retrieved item, when CMR2 retrieves a prior-list intrusion it updates context with the retrieval of that intrusion. This helps to explain why recall of an intrusion is often followed by recall termination (Miller, Kahana, & Weidemann, 2012), as it is more difficult to retrieve items associated to the current-list context. This also leads to two new predictions of the CMR2 model regarding successive recall of two PLIs. First, these two PLIs are more likely to be from neighboring lists, and second that PLIs recalled successively from the same list tend to have been presented in neighboring serial positions. In a meta-analysis of free recall studies, we found strong empirical support for these predictions (see Figure 3). CMR2 predicts these across-list contiguity effects for the same reason that it predicts within-list contiguity: Recall of an item reinstates its associated context, increasing the probability that a neighboring item is recalled next.⁶

In a new experiment, we tested a specific assumption of CMR2 regarding memory errors. Because all experimental items are stored in its memory, CMR2 will occasionally produce PLIs. To "suppress" these intrusions, CMR2 first retrieves an item based on the context cue, and then compares that item's associated context to the current context. PLIs generally are associated with contexts different from the end-of-list context used to initiate recall, and thus are more likely to be suppressed. To assess whether subjects perform recall in this way, we used the externalized free recall

procedure (EFR; Kahana et al., 2005; Unsworth & Brewer, 2010; Unsworth et al., 2010; Zaromb et al., 2006). In EFR, subjects verbalized all words that came to mind during recall and "rejected" any word that they thought was an error, indicated by a spacebar press immediately after that item. In CMR2 terms, this means that a subject would verbalize all retrieved items, and would reject an item when its associated context is dissimilar to the current context.

Using a single set of model parameters, CMR2 provided a good fit to both standard recall and EFR data. For both tasks, CMR2 captured the within-list recency of the probability of first recall and serial position curves, as well as recall of PLIs (see Figure 4). In EFR, CMR2 predicts that the probability of rejecting an item would be inversely related to its similarity to the current state of context. Thus, CMR2 accurately predicts that correct items would be rejected much less frequently than PLIs (Table 3; Kahana et al., 2005; Unsworth & Brewer, 2010; Unsworth et al., 2010; Zaromb et al., 2006). CMR2 further predicts the experimental finding that PLIs of greater list lags would be rejected more frequently (Unsworth et al., 2010). Thus, this set of results provides strong evidence that subjects retrieve more items than they recall, and filter their recalls based on context similarity. One concern, however, is that subjects may only use this approach when they have access to the appropriate context cue, as in immediate free recall, but not when they must actively retrieve the relevant context, as is often the case outside of the laboratory.

The ability of subjects to effectively target items from lists studied before the most recent list has been proposed as a serious challenge to retrieved temporal context models (Usher et al., 2008). The argument is that retrieved context models lack an explicit "list-tagging" mechanism, making it unclear how they can distinguish neighboring items from different lists (e.g., distinguishing the end of List 1 from the beginning of List 2). In our simulations of the list-before-last paradigm, we showed how the context-comparison mechanism used to target current-list items could also be effectively used to selectively target recall of items studied on a specific prior list. We showed that CMR2 can focus recall to the list before the last (Shiffrin, 1970) without a list-tagging mechanism by assuming that there is a shift in context between lists. This allows items from each list to be associated with a specific set of temporal contexts, thereby making them more easily accessible.

In our list-before-last simulations, we found that increasing the length of the intervening list reduced recall of target list items (a retroactive interference effect) but only when there was no recall between presentation of the target list ($n - 1$) and presentation of the intervening list (n). When recall was interpolated between lists, no reliable interference effect was observed. This pattern of results

⁶ Although our PLI contiguity effect could arguably have been induced by temporal autocorrelation in goodness of encoding across items, this hypothesis has been rejected in previous studies of contiguity in both free recall Howard, Youker, and Venkatadass (2008) and item recognition Schwartz, Howard, Jing, and Kahana (2005). Although the rarity of PLIs precludes our carrying out similar permutation tests here, the previous analyses of both Howard, Youker, and Venkatadass (2008) and Schwartz, Howard, Jing, and Kahana (2005) suggest that the contribution of the autocorrelation in goodness of encoding (perhaps representing attentional fluctuations) is small relative to the contribution of retrieved context or interitem associations, as predicted by CMR2.

is fully consistent with experimental data, as reported by Jang and Huber (2008) and Ward and Tan (2004). This paradigm was originally presented as evidence against classic decay and interference theories of forgetting: With recall between lists, increasing the length of the intervening list did not affect recall of target list items (Jang & Huber, 2008; Shiffrin, 1970; Ward & Tan, 2004). CMR2 also correctly predicts a greater proportion of intervening list intrusions when there is recall compared with no recall between lists (see Figure 7). In summary, CMR2 explains the ability to retrieve memories associated with a particular list by retrieving and distinguishing appropriate temporal contexts.

CMR2 shares several mechanisms with Jang and Huber (2008)'s model of the list-before-last paradigm. These include the idea that memories accumulate across lists and the idea that recalling an item causes a shift in the state of mental context. The assumption in CMR2 that proved most important in accounting for the list-before-last data is the idea that not all sampled items are recalled. This basic mechanism, which was absent in earlier retrieved context models, is needed to account for the fact that some of the items sampled during the initial retrieval phase are later recalled. Whereas in CMR any retrieved item was necessarily correct, this is not the case for CMR2.

Incorporating a generate-recognize mechanism is also a key component of the Jang and Huber (2008) model, and in CMR2 this mechanism is critical to reinstating the list-before-last context. By filtering out retrieved items based on their overlap with the current context, CMR2 can retrieve items from an earlier target list even though recall initiates with the current (end-of-list) context. In everyday life, recalling events from earlier contexts is usually aided by a multiplicity of other cues. For instance, to retrieve a memory of last Thanksgiving's dinner, probing one's memory with the venue of the dinner (Grandma's house) and relevant semantic features (e.g., turkey, cranberry sauce) may facilitate retrieval of the event. In the list-before-last paradigm, these other sources of organization are not provided and subjects must rely solely on temporal information. This is one reason why this task is so difficult, and also why it poses a serious challenge to models that lack explicit mechanisms for targeting memories based on temporal information.

In the list-before-last paradigm, CMR2 begins each recall period by covertly retrieving intervening-list intrusions, whose contexts are similar to the current context. However, once the first putative target-list item is recalled, CMR2 then attempts to recall other items with contexts similar to that item, which occasionally may be intervening-list items. Because such errors are more likely to occur after target items are recalled, CMR2 makes the novel prediction that the proportion of intervening-list intrusions recalled would increase with output position. We showed that this prediction is borne out in the experimental data. This may also relate to the finding in standard free recall that subjects are more likely to commit errors at later output positions (Kimball, Smith, & Kahana, 2007; Roediger & McDermott, 1995). For both types of recall instructions, as the number of retrieved contexts increases so does the probability that CMR2 retrieves a context from an incorrect list.

Lastly, we demonstrated that, by incorporating longstanding semantic associations into memory, CMR2 predicts the proactive interference caused by similar prior list items on recall of current list items (Loess, 1967; Wickens, 1970). Specifically, subjects exhibit markedly

increased difficulty in recalling list items when those items are from the same category as the words on the preceding list or lists. Switching to a new semantic category on a subsequent list "releases" subjects from this proactive interference effect, enabling them to recall more list items and commit fewer intrusions than on the same-category lists. CMR2 naturally predicts this pattern of results because of the dual effects of semantic and temporal similarity on recall. Recent prior list items become a powerful source of proactive interference when they are semantically similar to items on the target list. Reducing semantic similarity between lists enables the model to effectively target list items based on their strong associations to temporal context, thus, producing the "release from PI" effect.

Comparison to Other Models of Free Recall

CMR2 assumes that temporal context serves as the foundation for memory representations, organization, and search processes in free recall. In particular, CMR2 assumes that a single temporal context is updated with each studied or retrieved item, and that this implicitly serves to organize memories across lists. Several successful models of free recall assume that there is a separate list context, i.e., a context representation that changes only at the start of each list (e.g., Anderson & Bower, 1972; Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, & Usher, 2005; Kimball et al., 2007; Lehman & Malmberg, 2013; Raaijmakers & Shiffrin, 1981; Shiffrin & Steyvers, 1997; Sirotnin et al., 2005). These models also assume that within-list recall effects are primarily driven by item rather than context representations. Whereas these models may predict across-list recency effects based on list context, these models account for within-list recency effects by assuming the existence of a STM buffer. Similarly, the model introduced in Farrell (2012) assumes that items are grouped together during encoding, and that context representations are formed for each item, each group, and each list. Several of these models have been shown to account for some of the basic interlist effects; however, they have not been fit to the wide range of interlist phenomena examined here.

To account for interlist effects in free recall, CMR2 assumes that context changes with each presented, retrieved and distractor item. An alternative approach is to assume that different context features change at different temporal rates, directly encoding a representation of the time of occurrence of each item. This approach was taken by Howard, Shankar, Aue, and Criss (2015), who showed how such a scale invariant representation of temporal context could simultaneously account for both within-list and across-list recency and contiguity effects (see also Shankar & Howard, 2012).

Cognitive Control of Contextual Drift

In the list-before-last paradigm, time-of-test context strongly cues the intervening list but only weakly cues the target list. To effectively reinstate the target list context, CMR2 must first retrieve a target list item. In fitting the complex pattern of experimental data from this paradigm, CMR2 took on a high rate of context drift during recall as this would promote a more complete reinstatement of target list contexts following retrieval of a target list item. Comparing the best-fitting model parameters in immediate free recall with those in the list-before last paradigm provides further support for a changing role of context drift in the list-before-last paradigm. For the recall-between-lists condition, the rate of context drift between lists needed

to be much lower than in standard free recall (see Table 2). Whereas in free recall subjects may wish to forget all items from previous lists, in the list-before-last paradigm it is crucial to maintain strong representations of items from the previous list. Furthermore, the best-fit parameter for the context shift between lists was higher when there was a pause, rather than a recall period, between lists. This is consistent with the finding that the recall period itself may serve to distinguish successively presented lists (Pastötter et al., 2011; Szpunar et al., 2008; Tulving & Watkins, 1974), and that in CMR2 simulations of the list-before-last paradigm, recall distinguishes lists because of context updating from retrieved items. These results suggest that subjects can manipulate their internal context to improve access to those memories associated with the desired temporal context.

Studies of the directed forgetting paradigm, in which subjects are instructed to forget a list of just-presented items, provide some support for the notion that subjects can exert control over their internal context drift rate based on task demands. Specifically, the worsened memory of to-be-forgotten items as well as the improved memory of subsequently presented items can be explained by the theory that a subject induces a greater shift in context after the forget instruction (Lehman & Malmberg, 2009; Pastötter & Bäuml, 2007; Sahakyan & Delaney, 2003; Sahakyan & Kelley, 2002). CMR2 could be extended to examine the role of context change in directed forgetting by assuming there is a higher context drift rate immediately after a list that the subject is then instructed to forget. The CMR2 framework also suggests that subjects can interpret these context states to recall items from a particular context, as summarized in the next section.

Generate-Recognize Theory

Generate-recognize theory posits that recall involves two stages: generation and recognition. In response to a cue, one first generates candidate responses and then applies a recognition test to exclude candidate responses that do not exceed a familiarity or memory strength criterion (Bahrick, 1970; Postman, 1976). The recognition stage serves to exclude the recall of generated items that were not studied on the most recent (target) list. Subjects' ability to use this recognition mechanism would lead to lower recall of prior list items that were strongly associated with the retrieval cue. In support of this idea, findings obtained using the externalized free recall paradigm indicate that subjects generate more errors than they actually report (Kahana et al., 2005; Unsworth & Brewer, 2010; Unsworth et al., 2010; Zaromb et al., 2006).

CMR2 implements generate-recognize theory by restricting recall to retrieved items with contexts similar to the current context. The addition of this mechanism was one of the major modifications in extending CMR to CMR2. That CMR did not need this generate-recognize mechanism highlights one benefit of our current investigation: Requiring the model to have a more expansive memory lexicon revealed a mechanism necessary for explaining interlist effects. Although this mechanism is new to CMR2, a similar approach has been used in other models of free recall (Anderson & Bower, 1972; Malmberg & Shiffrin, 2005; Metcalfe, 1991; Raaijmakers & Shiffrin, 1981).

Whereas previous models have used a generate-recognize mechanism to filter out PLIs, CMR2 can also use this mechanism to target retrieval to a more distant list. Just as subjects may recognize an item to be from the most recent list based on its recency, they may also use this recency information to recall items from the list before the last.

This is consistent with findings that subjects are relatively accurate at judging how recently an item was presented, and may base such judgments on how strongly items are activated in memory (Hintzman, 2005).

Because of the item-driven dynamics of context in CMR2, the context-comparison process is inherently noisy: Items other than those from the list of interest can enter the competition for retrieval (see Equation A7). If an item from another list is retrieved, its context is incorporated into the cue used to probe memory, which further hinders performance (Zaromb et al., 2006). This makes CMR2 both more susceptible to recall incorrect items, as well as fail to recall correct items. We explore this latter type of memory error in the following section.

Failures in Recall

Tulving and Pearlstone (1966) distinguished between two kinds of retrieval failures: those that reflect a failure in availability (i.e., the item was not encoded effectively in memory) and those that reflect a failure in accessibility (i.e., the item was encoded effectively but cannot be retrieved because of the lack of an effective set of retrieval cues). That recall of categorized lists was higher when providing subjects with appropriate cues implies that decreased recall reflects a loss of accessibility rather than availability (Tulving & Pearlstone, 1966; Tulving & Psotka, 1971). These results underscore the importance of being able to access appropriate retrieval cues.

Our present investigation suggests that the context cue determines the accessibility of memories. Items presented more recently (whether more recently within a list or in a more recent list) are more likely to be recalled because they have relatively stronger associations to context. In a complementary way, the list-before-last task poses a challenge to the model because the appropriate context is not readily accessible but rather must be retrieved.

Accessibility of appropriate cues can also be used to explain the contiguity effects exhibited within and across lists: Recall of an item retrieves its associated contexts, thus increasing the context strengths of neighboring items. Although retrieved context models generally implicate context as the major determinant of recall dynamics, CMR2 demonstrates that the context cue can be used to target recalls to the desired list at the exclusion of other pre-existing or prior experimental items that may exist in memory.

When viewed in this way, PLIs are interpreted as items that benefit from easily accessible context cues. This was the case with the build-up of PI in Simulation 4, where PLIs benefited from strong semantic associations to items in the current list. It was also suggested in Simulation 1 that PLIs may benefit from stronger representations in context because of strong temporal or semantic associations to recently retrieved items. Thus, the same organizational factors that benefit recall of current list items may also increase recall of intrusions, leading to interference effects (Schacter, 1999; Schacter, Guerin, & St. Jacques, 2011). Such trade-offs between correct and incorrect recalls can only be examined in a framework in which memory is continuous across lists, and thus the memory system must use context and other cues to target retrieval of the appropriate items.

CMR2 embodies the view that much of the variability in memory search arises because of the changing nature of the retrieval cues present at the time of test (i.e., accessibility). However, variability in the availability of items, as defined in the model by contextual dynamics during encoding and by variability in associative represen-

tations updated during memory encoding, also undoubtedly plays an important role in memory retrieval. Although our current implementation of CMR2 does not fully represent these sources of variability during encoding, such extensions of the model are an important target for future work. CMR2 does assume that encoding varies with list position, such that primacy list items benefit from stronger associations to context. We further discuss the primacy effect in the next section.

CMR2 and the Primacy Effect

CMR2 predicts the primacy effect by assuming that early list items are more strongly associated to context. CMR2's implementation of the primacy effect could be developed to reflect cognitive factors benefiting primacy items. For instance, when subjects are instructed to overtly report their rehearsals, early list items benefit from the most rehearsals, thus improving the availability of those items during recall (Brodie & Murdock, 1977; Laming, 2008; Rundus, 1971; Tan & Ward, 2000). Further, Sederberg et al. (2006); Serruya et al. (2014) found that neural correlates of successful memory formation and attention were generally highest during presentation of the first list item, and decreased monotonically thereafter. Incorporating these processes into CMR2 could help to discern the contributions of each (for further discussion see Polyn et al., 2009).

Recently, Healey and Kahana (2014) showed that a simplified version of CMR2, CMR, can account for individual differences in the primacy effect, manifested both in recall probability and in recall initiation. Although 81 out of 126 subjects consistently initiated immediate free recall with the final list item (consistent with CMR2 predictions), a small but significant minority of subjects (34) were more likely to initiate recall with a primacy item (Positions 1–3) than the final list item, even after considerable practice. Healey and Kahana (2014) showed that CMR can account for both groups of subjects by altering the parameters governing the primacy gradient in the model (Equation A5). Although recency is a powerful cue for end of list items, early items are more strongly associated with context, and for some subjects the primacy items are more strongly associated to context than recency items, leading to initiation of recall with early list items. Although we do not consider individual differences here, CMR2 could just as easily consider variations in primacy to account for individual differences. Nonetheless, there are advantages to our approach of fitting the subject means, as CMR2 helps to inform our understanding of the contributions of a single context representation in memory search, which we outline below.

The Explanatory Scope of a Single Temporal Context Representation

Sederberg et al. (2008)'s version of the temporal context model (TCM-A) can explain within-list effects in both immediate free recall and continual distractor free recall. For the latter task, subjects perform a distractor task between each presented item, as well as after the final item before the recall period (Bjork & Whitten, 1974; Howard & Kahana, 1999). TCM-A predicts that contiguity and recency are maintained in continual-distractor free recall because items' relative strengths in context are functionally similar to those in immediate free recall, and the data support TCM-A's predictions. Here, we extend results of Sederberg et al. (2008) by showing that a single temporal context representation can account for contiguity and recency across

lists. Although it is common for a multilist model to assume that some aspect of the memory representation changes only between lists (Anderson & Bower, 1972; Farrell, 2012; Lehman & Malmberg, 2009; Sederberg et al., 2011; Sirotin et al., 2005), CMR2 predicts critical findings in standard free recall, as well as the more challenging list-before-last paradigm, simply by comparing the context of a retrieved item to the current state of context. Further, CMR2 can account for the influence of longstanding semantic representations on interference effects across lists.

We measure the success of the single-context assumption not only with the accuracy of CMR2's predictions, but also in the ability of the CMR2 model to explain how context retrieval mediates the results. To date, existing models of the list-before-last paradigm do not address how the list before last (target list) is reinstated after presentation of the intervening list, but rather assume that the probability of reinstatement changes based on the experimental manipulation (Jang & Huber, 2008; Lehman & Malmberg, 2009). CMR2 elaborates on this modeling work by suggesting how context retrieval may govern such probabilities. Specifically, the influence of intervening-list intrusions is mitigated by a task requiring context retrieval. In the case of the standard list-before-last paradigm, with CMR2's memory restricted to items presented in the experiment, this context retrieval reflects items from previous lists, including but not limited to the target list. More generally, tasks requiring context retrieval have been shown to produce a null list-length effect, even if the retrieved context is not associated with items presented previously in the experiment. For instance, in Experiment 3 of Jang and Huber (2008), a letter-completion task between lists effectively nullified the influence of intervening list-length on target-list recalls. Although CMR2 can only simulate free recall, the fact that the retrieval of nontarget list items can mitigate the influence of intervening-list length suggests that it is context retrieval in general, rather retrieval of target list contexts, which drives the effect of intervening list-length. CMR2's assumptions are also consistent with recent work by Divis and Benjamin (2014), who showed that in free recall, performing a semantic retrieval task versus an arithmetic task between lists enhanced subjects' recall for later lists and attenuated proactive interference from previous lists, and that these results are predicted by a model that assumes context fluctuates more quickly for the semantic than the arithmetic task.

CMR2 also informs the role of context retrieval in the list-before-last paradigm because it assumes that this retrieval process has an all-or-none effect on the subsequent recall period. Meaning, CMR2 predicts that probability of recall for target-list items does not vary as a function of the number of retrieved items in the previous recall period (so long as at least one item is retrieved). This prediction is consistent with the finding of Sahakyan and Hendricks (2012) that target-list recalls do not change with the level of retrieval in the previous recall period.

It is important to note that our results do not rule out the hypothesis that subjects have multiple representations of time, but rather suggest that a cognitive representation of a shorter time scale (in CMR2 terms, items) also provides the structure to organize items on a longer time scale (in CMR2 terms, lists). Indeed, recent studies of neural activity suggest that this is the case. In a series of studies, Hasson and colleagues presented subjects with movie or audio clips that were scrambled at different time scales. (For instance, in one experiment they divided a 4-min movie clip into 4 s, 12 s, or 36 s segments, and

randomly shuffled those segments.) They examined the reliability of brain activity as a function of time scale, and found that early sensory regions (e.g., primary visual or auditory cortex) exhibited high reliability at all time scales, yet higher order areas (e.g., precuneus) only exhibited high reliability for clips that were scrambled with longer segments (Hasson, Yang, Vallines, Heeger, & Rubin, 2008; Honey, Thesen, Donner, Silbert, Carlson, Devinsky, & Hasson, 2012; Lerner, Honey, Silbert, & Hasson, 2011). This suggests that lower sensory regions process information at shorter time scales and provide this information to higher order areas, which in turn accumulate information at longer time scales. Significantly, Honey et al. (2012) measured electrocorticographic oscillatory power to characterize the precise time scale of brain activity giving rise to this relationship. They found that, whereas fast oscillatory activity correlated with short segments, slow oscillatory activity correlated with long segments, thus suggesting that the former informs the latter. CMR2 proposes a similar approach, in that the cognitive representations at shorter time scales influences the representations at longer time scales.

In the memory laboratory, the complexity of human experience is reduced to sequences of items occurring at predetermined times, often organized into groupings that we call lists. Thus, it is not uncommon for computational models of list learning to assume that there are distinct representations of items and lists. However, outside of the laboratory, the continuous stream of experiences lack such a well-defined structure, and it is not as straightforward to define or distinguish the different time scales that influence memory. As an alternate approach to making assumptions about what the real-world (or experimental) time scales might be, we presented simulations of a computational memory model that assumes temporal context is updated only on a single time scale (items) and that this representation serves to organize memory on other time scales (e.g., lists). As such, CMR2 can be used to test the limits of using temporal contexts to guide memory search. Our CMR2 simulations demonstrated how a temporal context defined by items suffices as a context to distinguish lists, which raises the possibility (supported by neural findings) that a representation of temporal context that evolves on the basis of items can account for memory both at the level of items and lists. The question of whether an item based context model, such as CMR2, can account for the complexities of real world memory across multiple time scales awaits future work.

Conclusion

Our development of the CMR2 model enabled us to extend the explanatory scope of retrieved context theory beyond the domain of within-list recall phenomena such as recency and primacy effects, and temporal and semantic clustering effects. We tested CMR2 on its ability to account for both proactive and retroactive interference effects, including the release from proactive interference resulting from a change in semantic category across subsequent lists. We also evaluated the model's ability to account for data on recall errors (intrusions) both when recalling the most recent list, as is typically done in laboratory experiments, and when trying to recall an earlier target list. A model with a single contextual drift mechanism worked surprisingly well to simultaneously account for both within-list and across-list memory effects. Although the success of this model cannot rule out the possibility of hierarchical context representations, it suggests that such representations can be viewed as varying along a continuum rather than as discrete levels. Our analyses of CMR2 also

highlighted the importance of modeling a post-retrieval editing mechanism in any description of human recall performance. Although early theoretical work has emphasized the importance of such a mechanism (Atkinson & Juola, 1974; Bahrick, 1970; Kintsch, 1970; Postman, 1976), theorists have frequently overlooked this process when applying their models to single list recall paradigms. In summary, expanding the memory of a retrieved context model of free recall revealed several significant components of memory search, many of which did not play as critical a role when considering within list effects alone. The development of memory theories and models will continue to illuminate the complexity of memory dynamics.

References

- Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review*, *79*, 97–123. <http://dx.doi.org/10.1037/h0033773>
- Aslan, A., & Bäuml, K.-H. (2008). Memorial consequences of imagination in children and adults. *Psychonomic Bulletin & Review*, *15*, 833–837. <http://dx.doi.org/10.3758/PBR.15.4.833>
- Atkinson, R. C., & Juola, J. F. (1974). Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology*. San Francisco, CA: Freeman.
- Bahrick, H. P. (1970). Two-phase model for prompted recall. *Psychological Review*, *77*, 215–222. <http://dx.doi.org/10.1037/h0029099>
- Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology*, *6*, 173–189. [http://dx.doi.org/10.1016/0010-0285\(74\)90009-7](http://dx.doi.org/10.1016/0010-0285(74)90009-7)
- Bower, G. H. (1967). A multicomponent theory of the memory trace. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 1, pp. 229–325). New York, NY: Academic Press.
- Bridge, D. (2006). *Memory and cognition: What difference does gender make?* Unpublished Honor's thesis, Syracuse University.
- Brodie, D. A., & Murdock, B. B. (1977). Effects of presentation time on nominal and functional serial position curves in free recall. *Journal of Verbal Learning and Verbal Behavior*, *16*, 185–200. [http://dx.doi.org/10.1016/S0022-5371\(77\)80046-7](http://dx.doi.org/10.1016/S0022-5371(77)80046-7)
- Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review*, *107*, 127–181. <http://dx.doi.org/10.1037/0033-295X.107.1.127>
- Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, *106*, 551–581. <http://dx.doi.org/10.1037/0033-295X.106.3.551>
- 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. <http://dx.doi.org/10.1037/0033-295X.112.1.3>
- Davelaar, E. J., Usher, M., Haarmann, H. J., & Goshen-Gottstein, Y. (2008). Postscript: Through TCM, STM shines bright. *Psychological Review*, *115*, 1116–1118. <http://dx.doi.org/10.1037/0033-295X.115.4.1116>
- Dennis, S., & Humphreys, M. S. (2001). A context noise model of episodic word recognition. *Psychological Review*, *108*, 452–478. <http://dx.doi.org/10.1037/0033-295X.108.2.452>
- Divis, K., & Benjamin, A. (2014). Retrieval speeds context fluctuation: Why semantic generation enhances later learning but hinders prior learning. *Memory & Cognition*, *42*, 1049–1062. <http://dx.doi.org/10.3758/s13421-014-0425-y>
- Duncan, M., & Lewandowsky, S. (2005). The time course of response suppression: No evidence for a gradual release from inhibition. *Memory*, *13*, 236–246. <http://dx.doi.org/10.1080/09658210344000233>

- Estes, W. K. (1955a). Statistical theory of distributional phenomena in learning. *Psychological Review*, *62*, 369–377. <http://dx.doi.org/10.1037/h0046888>
- Estes, W. K. (1955b). Statistical theory of spontaneous recovery and regression. *Psychological Review*, *62*, 145–154. <http://dx.doi.org/10.1037/h0048509>
- Farrell, S. (2012). Temporal clustering and sequencing in short-term memory and episodic memory. *Psychological Review*, *119*, 223–271. <http://dx.doi.org/10.1037/a0027371>
- Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review*, *9*, 59–85. <http://dx.doi.org/10.3758/BF03196257>
- Farrell, S., & Lewandowsky, S. (2008). Empirical and theoretical limits on lag recency in free recall. *Psychonomic Bulletin & Review*, *15*, 1236–1250. <http://dx.doi.org/10.3758/PBR.15.6.1236>
- Golomb, J. D., Peelle, J. E., Addis, K. M., Kahana, M. J., & Wingfield, A. (2008). Effects of adult aging on utilization of temporal and semantic associations during free and serial recall. *Memory & Cognition*, *36*, 947–956. <http://dx.doi.org/10.3758/MC.36.5.947>
- Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A hierarchy of temporal receptive windows in human cortex. *Journal of Neuroscience*, *28*, 2539–2550. <http://dx.doi.org/10.1523/JNEUROSCI.5487-07.2008>
- Healey, M. K., & Kahana, M. J. (2014). Is memory search governed by universal principles or idiosyncratic strategies? *Journal of Experimental Psychology: General*, *143*, 575–596. <http://dx.doi.org/10.1037/a0033715>
- Henson, R. (1998). Short-term memory for serial order: The start-end model. *Cognitive Psychology*, *36*, 73–137. <http://dx.doi.org/10.1006/cogp.1998.0685>
- Hintzman, D. L. (2005). Memory strength and recency judgments. *Psychonomic Bulletin & Review*, *12*, 858–864. <http://dx.doi.org/10.3758/BF03196777>
- Hintzman, D. L., & Block, R. A. (1971). Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology*, *88*, 297–306. <http://dx.doi.org/10.1037/h0030907>
- Honey, C. J., Thesen, T., Donner, T. H., Silbert, L. J., Carlson, C. E., Devinsky, O., . . . Hasson, U. (2012). Slow cortical dynamics and the accumulation of information over long timescales. *Neuron*, *76*, 423–434. <http://dx.doi.org/10.1016/j.neuron.2012.08.011>
- Howard, M. W. (2004). Scaling behavior in the temporal context model. *Journal of Mathematical Psychology*, *48*, 230–238. <http://dx.doi.org/10.1016/j.jmp.2004.03.004>
- 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. <http://dx.doi.org/10.1037/0278-7393.25.4.923>
- Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, *46*, 269–299. <http://dx.doi.org/10.1006/jmps.2001.1388>
- Howard, M. W., Kahana, M. J., & Wingfield, A. (2006). Aging and contextual binding: Modeling recency and lag-recency effects with the temporal context model. *Psychonomic Bulletin & Review*, *13*, 439–445. <http://dx.doi.org/10.3758/BF03193867>
- Howard, M. W., Shankar, K. H., Aue, W. R., & Criss, A. H. (2015). A distributed representation of internal time. *Psychological Review*, *122*, 24–53. <http://dx.doi.org/10.1037/a0037840>
- Howard, M. W., Venkatadass, V., Norman, K. A., & Kahana, M. J. (2007). Associative processes in immediate recency. *Memory & Cognition*, *35*, 1700–1711. <http://dx.doi.org/10.3758/BF03193503>
- Howard, M. W., Youker, T. E., & Venkatadass, V. (2008). The persistence of memory: Contiguity effects across hundreds of seconds. *Psychonomic Bulletin & Review*, *15*, 58–63. <http://dx.doi.org/10.3758/PBR.15.1.58>
- Jacoby, L. L., & Hollingshead, A. (1990). Toward a generate/recognize model of performance on direct and indirect tests of memory. *Journal of Memory and Language*, *29*, 433–454. [http://dx.doi.org/10.1016/0749-596X\(90\)90065-8](http://dx.doi.org/10.1016/0749-596X(90)90065-8)
- Jang, Y., & Huber, D. (2008). Context retrieval and context change in free recall: Recalling from long-term memory drives list isolation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 112–127. <http://dx.doi.org/10.1037/0278-7393.34.1.112>
- Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, *24*, 103–109. <http://dx.doi.org/10.3758/BF03197276>
- Kahana, M. J. (2012). *Foundations of human memory*. New York, NY: Oxford University Press.
- Kahana, M. J., Dolan, E. D., Sauder, C. L., & Wingfield, A. (2005). Intrusions in episodic recall: Age differences in editing of overt responses. *Journal of Gerontology: Psychological Sciences*, *60*, 92–97. <http://dx.doi.org/10.1093/geronb/60.2.P92>
- Kahana, M. J., Howard, M. W., Zaremb, F., & Wingfield, A. (2002). Age dissociates recency and lag recency effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 530–540. <http://dx.doi.org/10.1037/0278-7393.28.3.530>
- Keppel, G. (1968). Retroactive and proactive inhibition. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory* (pp. 172–213). Englewood Cliffs, NJ: Prentice Hall.
- Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior*, *1*, 153–161. [http://dx.doi.org/10.1016/S0022-5371\(62\)80023-1](http://dx.doi.org/10.1016/S0022-5371(62)80023-1)
- Kimball, D. R., Smith, T. A., & Kahana, M. J. (2007). The fSAM model of false recall. *Psychological Review*, *114*, 954–993. <http://dx.doi.org/10.1037/0033-295X.114.4.954>
- Kintsch, (1970). *Learning, memory, and conceptual processes*. New York, NY: Wiley.
- Laming, D. (2008). An improved algorithm for predicting free recalls. *Cognitive Psychology*, *57*, 179–219. <http://dx.doi.org/10.1016/j.cogpsych.2008.01.001>
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211–240. <http://dx.doi.org/10.1037/0033-295X.104.2.211>
- Lehman, M., & Malmberg, K. J. (2009). A global theory of remembering and forgetting from multiple lists. *Journal Experimental Psychology: Learning, Memory, and Cognition*, *35*, 970–988. <http://dx.doi.org/10.1037/a0015728>
- Lehman, M., & Malmberg, K. J. (2013). A buffer model of memory encoding and temporal correlations in retrieval. *Psychological Review*, *120*, 155–189. <http://dx.doi.org/10.1037/a0030851>
- Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *Journal of Neuroscience*, *31*, 2906–2915. <http://dx.doi.org/10.1523/JNEUROSCI.3684-10.2011>
- Levy, W. B., & Steward, O. (1983). Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. *Neuroscience*, *8*, 791–797. [http://dx.doi.org/10.1016/0306-4522\(83\)90010-6](http://dx.doi.org/10.1016/0306-4522(83)90010-6)
- Lewandowsky, S. (1999). Redintegration and response suppression in serial recall: A dynamic network model. *International Journal of Psychology*, *34*, 434–446. <http://dx.doi.org/10.1080/002075999399792>
- Lewandowsky, S., & Farrell, S. (2008). Short-term memory: New data and a model. *The Psychology of Learning and Motivation*, *49*, 1–48. [http://dx.doi.org/10.1016/S0079-7421\(08\)00001-7](http://dx.doi.org/10.1016/S0079-7421(08)00001-7)
- Lewandowsky, S., & Murdock, B. B. (1989). Memory for serial order. *Psychological Review*, *96*, 25–57. <http://dx.doi.org/10.1037/0033-295X.96.1.25>
- Loess, H. (1967). Short-term memory, word class, and sequence of items. *Journal of Experimental Psychology*, *74*, 556–561. <http://dx.doi.org/10.1037/h0024787>

- Lohnas, L. J., & Kahana, M. J. (2013). Parametric effects of word frequency effect in memory for mixed frequency lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1943–1946. <http://dx.doi.org/10.1037/a0033669>
- Lohnas, L. J., & Kahana, M. J. (2014). Compound cuing in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 12–24. <http://dx.doi.org/10.1037/a0033698>
- Lohnas, L. J., Polyn, S. M., & Kahana, M. J. (2011). Contextual variability in free recall. *Journal of Memory and Language*, *64*, 249–255. <http://dx.doi.org/10.1016/j.jml.2010.11.003>
- Malmberg, K. J., & Shiffrin, R. M. (2005). The “one-shot” hypothesis for context storage. *Journal Experimental Psychology: Learning, Memory, and Cognition*, *31*, 322–336. <http://dx.doi.org/10.1037/0278-7393.31.2.322>
- McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychological Review*, *39*, 352–370. <http://dx.doi.org/10.1037/h0069819>
- Melton, A. W., & Irwin, J. M. (1940). The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *American Journal of Psychology*, *53*, 173–203. <http://dx.doi.org/10.2307/1417415>
- Melton, A. W., & von Lackum, W. J. (1941). Retroactive and proactive inhibition in retention: Evidence for a two-factor theory of retroactive inhibition. *American Journal of Psychology*, *54*, 157–173. <http://dx.doi.org/10.2307/1416789>
- Mensink, G.-J. M., & Raaijmakers, J. G. W. (1988). A model for interference and forgetting. *Psychological Review*, *95*, 434–455. <http://dx.doi.org/10.1037/0033-295X.95.4.434>
- Metcalfe, J. (1991). Recognition failure and the composite memory trace in CHARM. *Psychological Review*, *98*, 529–553. <http://dx.doi.org/10.1037/0033-295X.98.4.529>
- Miller, J. F., Kahana, M. J., & Weidemann, C. T. (2012). Recall termination in free recall. *Memory & Cognition*, *40*, 540–550. <http://dx.doi.org/10.3758/s13421-011-0178-9>
- Murdock, B. B. (1961). The retention of individual items. *Journal of Experimental Psychology*, *62*, 618–625. <http://dx.doi.org/10.1037/h0043657>
- Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, *64*, 482–488. <http://dx.doi.org/10.1037/h0045106>
- Murdock, B. B. (1974). *Human memory: Theory and data*. Potomac, MD: Erlbaum and Associates.
- Murdock, B. B. (1997). Context and mediators in a theory of distributed associative memory (TODAM2). *Psychological Review*, *104*, 839–862. <http://dx.doi.org/10.1037/0033-295X.104.4.839>
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments and Computers*, *36*, 402–407. <http://dx.doi.org/10.3758/BF03195588>
- Page, M. P. A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review*, *105*, 761–781. <http://dx.doi.org/10.1037/0033-295X.105.4.761-781>
- Pastötter, B., & Bäuml, K.-H. (2007). The crucial role of post-cue encoding in directed forgetting and context-dependent forgetting. *Journal Experimental Psychology: Learning, Memory, and Cognition*, *33*, 977–982. <http://dx.doi.org/10.1037/0278-7393.33.5.977>
- Pastötter, B., Bäuml, K.-H., & Hanslmayr, S. (2008). Oscillatory brain activity before and after an internal context change: Evidence for a reset of encoding processes. *NeuroImage*, *43*, 173–181. <http://dx.doi.org/10.1016/j.neuroimage.2008.07.005>
- Pastötter, B., Schicker, S., Niedernhuber, J., & Bäuml, K.-H. T. (2011). Retrieval during learning facilitates subsequent memory encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 287–297. <http://dx.doi.org/10.1037/a0021801>
- 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. <http://dx.doi.org/10.1037/a0014420>
- Postman, L. (1976). Interference theory revisited. In J. Brown (Ed.), *Recall and recognition* (p. 157–181). Oxford, England: Wiley and Sons.
- Postman, L., & Hasher, L. (1972). Conditions of proactive inhibition in free recall. *Journal of Experimental Psychology*, *92*, 276–284. <http://dx.doi.org/10.1037/h0032074>
- Postman, L., & Underwood, B. J. (1973). Critical issues in interference theory. *Memory & Cognition*, *1*, 19–40. <http://dx.doi.org/10.3758/BF03198064>
- Raaijmakers, J. G. W., & Shiffrin, R. M. (1980). SAM: A theory of probabilistic search of associative memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 14, p. 207–262). New York, NY: Academic Press.
- Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, *88*, 93–134. <http://dx.doi.org/10.1037/0033-295X.88.2.93>
- Rao, V. A., & Howard, M. W. (2008). Retrieved context and the discovery of semantic structure. In J. Platt, D. Koller, Y. Singer, & S. Roweis (Eds.), *Advances in neural information processing systems* (p. 1193). MIT Press: Cambridge, MA.
- Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 803–814. <http://dx.doi.org/10.1037/0278-7393.21.4.803>
- Rundus, D. (1971). An analysis of rehearsal processes in free recall. *Journal of Experimental Psychology*, *89*, 63–77. <http://dx.doi.org/10.1037/h0031185>
- Sahakyan, L., & Delaney, P. F. (2003). Can encoding differences explain the benefits of directed forgetting in the list method paradigm? *Journal of Memory and Language*, *48*, 195–206. [http://dx.doi.org/10.1016/S0749-596X\(02\)00524-7](http://dx.doi.org/10.1016/S0749-596X(02)00524-7)
- Sahakyan, L., & Hendricks, H. E. (2012). Context change and retrieval difficulty in the list-before-last paradigm. *Memory & Cognition*, *40*, 844–860. <http://dx.doi.org/10.3758/s13421-012-0198-0>
- Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 1064–1072. <http://dx.doi.org/10.1037/0278-7393.28.6.1064>
- Schacter, D. L. (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist*, *54*, 182–203. <http://dx.doi.org/10.1037/0003-066X.54.3.182>
- Schacter, D. L., Guerin, S. A., & St. Jacques, P. L. (2011). Memory distortion: An adaptive perspective. *Trends in Cognitive Sciences*, *15*, 467–474. <http://dx.doi.org/10.1016/j.tics.2011.08.004>
- Schwartz, G., Howard, M. W., Jing, B., & Kahana, M. J. (2005). Shadows of the past: Temporal retrieval effects in recognition memory. *Psychological Science*, *16*, 898–904. <http://dx.doi.org/10.1111/j.1467-9280.2005.01634.x>
- Sederberg, P. B., Gauthier, L. V., Terushkin, V., Miller, J. F., Barnathan, J. A., & Kahana, M. J. (2006). Oscillatory correlates of the primacy effect in episodic memory. *NeuroImage*, *32*, 1422–1431. <http://dx.doi.org/10.1016/j.neuroimage.2006.04.223>
- Sederberg, P. B., Gershman, S. J., Polyn, S. M., & Norman, K. A. (2011). Human memory reconsolidation can be explained using the temporal context model. *Psychonomic Bulletin & Review*, *18*, 455–468. <http://dx.doi.org/10.3758/s13423-011-0086-9>
- 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. <http://dx.doi.org/10.1037/a0013396>
- Sederberg, P. B., Miller, J. F., Howard, W. H., & Kahana, M. J. (2010). The temporal contiguity effect predicts episodic memory performance.

- Memory & Cognition*, 38, 689–699. <http://dx.doi.org/10.3758/MC.38.6.689>
- Serruya, M. D., Sederberg, P. B., & Kahana, M. J. (2014). Power shifts track serial position and modulate encoding in human episodic memory. *Cerebral Cortex*, 24, 403–413. <http://dx.doi.org/10.1093/cercor/bhs318>
- Shankar, K. H., & Howard, M. W. (2012). A scale-invariant internal representation of time. *Neural Computation*, 24, 134–193. http://dx.doi.org/10.1162/NECO_a_00212
- Shiffrin, R. M. (1970). Forgetting: Trace erosion or retrieval failure? *Science*, 168, 1601–1603. <http://dx.doi.org/10.1126/science.168.3939.1601>
- Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM—Retrieving effectively from memory. *Psychonomic Bulletin and Review*, 4, 145. <http://dx.doi.org/10.3758/BF03209391>
- Sirotin, Y. B., Kimball, D. R., & Kahana, M. J. (2005). Going beyond a single list: Modeling the effects of prior experience on episodic free recall. *Psychonomic Bulletin & Review*, 12, 787–805. <http://dx.doi.org/10.3758/BF03196773>
- Steyvers, M., Shiffrin, R. M., & Nelson, D. L. (2004). Word association spaces for predicting semantic similarity effects in episodic memory. In A. F. Healy (Ed.), *Cognitive psychology and its applications: Festschrift in honor of Lyle Bourne, Walter Kintsch, and Thomas Landauer*. Washington, DC: American Psychological Association.
- Szpunar, K. K., McDermott, K. B., & Roediger, H. L. (2008). Testing during study insulates against the buildup of proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1392–1399. <http://dx.doi.org/10.1037/a0013082>
- 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. <http://dx.doi.org/10.1037/0278-7393.26.6.1589>
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). New York, NY: Academic Press.
- Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 5, 381–391. [http://dx.doi.org/10.1016/S0022-5371\(66\)80048-8](http://dx.doi.org/10.1016/S0022-5371(66)80048-8)
- Tulving, E., & Psotka, J. (1971). Retroactive inhibition in free recall: Inaccessibility of information available in the memory store. *Journal of Experimental Psychology*, 87, 1–8. <http://dx.doi.org/10.1037/h0030185>
- Tulving, E., & Rosenbaum, R. S. (2006). What do explanations of the distinctiveness effect need to explain? In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and memory* (pp. 407–423). New York, NY: Oxford University Press.
- Tulving, E., & Watkins, M. J. (1974). On negative transfer: Effects of testing one list on the recall of another. *Journal of Verbal Learning and Verbal Behavior*, 13, 181–193. [http://dx.doi.org/10.1016/S0022-5371\(74\)80043-5](http://dx.doi.org/10.1016/S0022-5371(74)80043-5)
- Underwood, B. J. (1945). The effect of successive interpolations and proactive inhibition. *Psychological Monographs*, 59, i–33. <http://dx.doi.org/10.1037/h0093547>
- Underwood, B. J. (1957). Interference and forgetting. *Psychological Review*, 64, 49–60. <http://dx.doi.org/10.1037/h0044616>
- Underwood, B. J. (1969). *Some correlates of item repetition in free-recall learning*, 8, 83–94. [http://dx.doi.org/10.1016/S0022-5371\(69\)80015-0](http://dx.doi.org/10.1016/S0022-5371(69)80015-0)
- Unsworth, N. (2008). Exploring the retrieval dynamics of delayed and final free recall: Further evidence for temporal-contextual search. *Journal of Memory and Language*, 59, 223–236. <http://dx.doi.org/10.1016/j.jml.2008.04.002>
- Unsworth, N., & Brewer, G. (2010). Variation in working memory capacity and intrusions: Differences in generation or editing? *European Journal of Cognitive Psychology*, 22, 990–1000. <http://dx.doi.org/10.1080/09541440903175086>
- Unsworth, N., Brewer, G., & Spillers, G. (2010). Understanding the dynamics of correct and error responses in free recall: Evidence from externalized free recall. *Memory & Cognition*, 38, 419. <http://dx.doi.org/10.3758/MC.38.4.419>
- Unsworth, N., Brewer, G. A., & Spillers, G. J. (2013). Focusing the search: Proactive and retroactive interference and the dynamics of free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 1742–1756. <http://dx.doi.org/10.1037/a0033743>
- Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, 114, 104–132. <http://dx.doi.org/10.1037/0033-295X.114.1.104>
- Unsworth, N., Spillers, G. J., & Brewer, G. A. (2012). Evidence for noisy contextual search: Examining the dynamics of list-before-last recall. *Memory*, 20, 1–13. <http://dx.doi.org/10.1080/09658211.2011.626430>
- Usher, M., Davelaar, E. J., Haarmann, H. J., & Goshen-Gottstein, Y. (2008). Short-term memory after all: Comment on Sederberg, Howard and Kahana (2008). *Psychological Review*, 115, 1108–1118. <http://dx.doi.org/10.1037/a0013725>
- Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, 108, 550–592. <http://dx.doi.org/10.1037/0033-295X.108.3.550>
- Ward, G., & Tan, L. (2004). The effect of the length of to-be-remembered lists and intervening lists on free recall: A reexamination using overt rehearsal. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 1196–1210. <http://dx.doi.org/10.1037/0278-7393.30.6.1196>
- Wickens, D. D. (1970). Encoding categories of words: An empirical approach to meaning. *Psychological Review*, 77, 1–15. <http://dx.doi.org/10.1037/h0028569>
- Wickens, D. D., Born, D. G., & Allen, C. K. (1963). Proactive inhibition and item similarity in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 2, 440–445. [http://dx.doi.org/10.1016/S0022-5371\(63\)80045-6](http://dx.doi.org/10.1016/S0022-5371(63)80045-6)
- Wixted, J. T., & Rohrer, D. (1994). Analyzing the dynamics of free recall: An integrative review of the empirical literature. *Psychonomic Bulletin and Review*, 1, 89–106. <http://dx.doi.org/10.3758/BF03200763>
- Yntema, D. B., & Trask, F. P. (1963). Recall as a search process. *Journal of Verbal Learning and Verbal Behavior*, 2, 65–74. [http://dx.doi.org/10.1016/S0022-5371\(63\)80069-9](http://dx.doi.org/10.1016/S0022-5371(63)80069-9)
- Zaromb, F. M., Howard, M. W., Dolan, E. D., Sirotin, Y. B., Tully, M., Wingfield, A., & Kahana, M. J. (2006). Temporal associations and prior-list intrusions in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 792–804. <http://dx.doi.org/10.1037/0278-7393.32.4.792>

(Appendices follow)

Appendix A

Formal Description of the Context Maintenance and Retrieval Model With Continuous Memory (CMR2)

Structure and Initialization

CMR2 has two representational layers, each defined as a vector: the feature layer F and the context layer C . The i^{th} item presented to the model activates its associated features in F (Bower, 1967; Underwood, 1969) and is denoted by the vector \mathbf{f}_i . This in turn updates the state of context to the vector \mathbf{c}_i , as described below. Each of these vectors is initialized to $\mathbf{0}$.

The layers interact through two associative matrices: M^{FC} , which stores the strengths of the feature-to-context associations, and M^{CF} , which stores the strength of the context-to-feature associations. Each association matrix is a weighted sum of a pre-experimental component (M_{pre}^{FC} and M_{pre}^{CF}) and an experimental (episodic) component (M_{exp}^{FC} and M_{exp}^{CF}). At the start of each experimental session, the experimental associations are initialized to 0. We make the simplifying assumption that M_{pre}^{FC} is initialized to an identity matrix. The semantic relations between items are implemented in M_{pre}^{CF} using Latent Semantic Analysis (LSA; Landauer & Dumais, 1997). LSA is a mathematical technique that decomposes large bodies of text into a multidimensional model of semantic space. The cosine of the angle between two words' vector representations in multidimensional feature space serves as the LSA measure of similarity. Thus, each element (a, b) in M_{pre}^{CF} is determined by taking the $\cos\theta$ similarity value between words a and b (though here we define an item's similarity with itself as 0).

Item Presentation

For simplicity we assume that each item has a localist representation, where \mathbf{f}_i is equal to 1 for the i^{th} element and is 0 for all other elements. The first item presented in each simulated session is not a list item, but rather a distraction item used to initialize the state of the context vector to have nonzero elements. Each \mathbf{f}_i is used to determine the input to C :

$$\mathbf{c}_i^{\text{IN}} = \frac{M^{FC}\mathbf{f}_i}{\|M^{FC}\mathbf{f}_i\|}. \quad (\text{A1})$$

The new context state integrates according to Equation 1. In this way, context is a recency-weighted sum of past context states. During item encoding, β determines how much new information (\mathbf{c}_i^{IN}) is added into context with each studied item. ρ_i weakens the previous state of context such that $\|\mathbf{c}_i\| = 1$:

$$\rho_i = \sqrt{1 + \beta^2[(\mathbf{c}_{i-1} \cdot \mathbf{c}_i^{\text{IN}})^2 - 1]} - \beta(\mathbf{c}_{i-1} \cdot \mathbf{c}_i^{\text{IN}}). \quad (\text{A2})$$

Forming Associations Between Items and Context

Each context state during study is used to update association matrices between item and context vectors using a Hebbian outer-product learning rule:

$$\begin{aligned} \Delta M_{exp}^{FC} &= \mathbf{c}_{i-1}\mathbf{f}_i^{\top} \\ \Delta M_{exp}^{CF} &= \mathbf{f}_i\mathbf{c}_{i-1}^{\top} \end{aligned} \quad (\text{A3})$$

Note that the first presented item, \mathbf{f}_1 , is associated with the first state of context, $\mathbf{c}_0 = \mathbf{0}$. Thus, this item is not incorporated into the association matrices.

We choose this association definition based on previously published versions of CMR (Howard et al., 2006; Sederberg et al., 2008) rather than the method presented in Polyn et al. (2009), as our definition is more physiologically plausible given the asymmetry of Hebbian associations (Levy & Steward, 1983). Because \mathbf{c}_i is derived from \mathbf{f}_i , associating these two states would require that \mathbf{f}_i is held in mind to later associate with \mathbf{c}_i . Thus, \mathbf{c}_{i-1} is a readily available cognitive state to associate with \mathbf{f}_i . Nonetheless, defining associations in this way, rather than the method of Polyn et al. (2009), does not affect the ability of CMR2 to predict the results reported here.

The relative strengths of the pre-experimental and experimental associations are controlled by parameters γ_{FC} , γ_{CF} :

$$\begin{aligned} M^{FC} &= (1 - \gamma_{FC})M_{pre}^{FC} + \gamma_{FC}M_{exp}^{FC}, \\ M^{CF} &= (1 - \gamma_{CF})(I + sM_{pre}^{CF}) + \gamma_{CF}\phi_i M_{exp}^{CF}. \end{aligned} \quad (\text{A4})$$

(Appendices continue)

γ_{FC} influences the magnitude of the tendency to make forward transitions during recall (Howard & Kahana, 2002). I is an identity matrix the same size as M_{pre}^{CF} . Effectively this means that the on-diagonal terms are not multiplied by the s parameter. This allows the s parameter to scale semantic relations between pairs of different items while having no effect on auto-associations.

ϕ_i scales the magnitude of context-to-feature connections to simulate increased attention to beginning-of-list items:

$$\phi_i = \phi_s e^{-\phi_d(i-1)} + 1, \tag{A5}$$

where ϕ_s is a model parameter that scales the overall level of primacy, and the model parameter ϕ_d scales the degree to which primacy decays for each list item presented after the first item. Because we only consider primacy effects inasmuch as they influence serial position effects, we could have chosen an attentional process without introducing two additional free parameters (e.g., Howard et al., 2006). However, because this form of the primacy gradient has been used in earlier versions of the model (Polyn et al., 2009; Sederberg et al., 2008), we kept this primacy mechanism in CMR2 to avoid introducing differences in model predictions based on our choice of primacy mechanism. In the General Discussion we present alternate possibilities for incorporating primacy mechanisms into CMR2.

The Recall Process

Once all items on a list have been presented, the time-of-test context is used to activate each item according to

$$\mathbf{f}_t^{\text{IN}} = M^{CF} \mathbf{c}_t, \tag{A6}$$

where \mathbf{f}_t^{IN} is a vector of activation values, one corresponding to each presented item. We then assign the l ($= 4 \times$ list-length) highest activation values to a vector \mathbf{a} , as it is extremely rare for items with the lowest activation values to be retrieved by the model, yet including such items is extremely computationally expensive for the process described below. The l activation values are used as the starting values for a leaky accumulator process (Usher & McClelland, 2001). The n^{th} step of this process is given by:

$$\begin{aligned} \mathbf{x}_n &= (1 - \tau\kappa - \tau\lambda\mathbf{N})\mathbf{x}_{n-1} + \tau\mathbf{a} + \boldsymbol{\epsilon}, \\ \mathbf{x}_n &\rightarrow \max(\mathbf{x}_n, \mathbf{0}). \end{aligned} \tag{A7}$$

Each element of the vector \mathbf{x}_n corresponds to an element in \mathbf{a} . Because $\mathbf{x}_0 = \mathbf{0}$, the activation for each item given in \mathbf{a} is used as its starting line in the race to threshold. τ is a fixed time constant, κ is a parameter that determines the decay rate for item activations,

and λ is the lateral inhibition parameter, scaling the strength of an inhibitory matrix \mathbf{N} that subtracts each item's activations from all of the others except itself. $\boldsymbol{\epsilon}$ is a random vector whose elements are drawn from a random normal distribution with mean 0 and SD as a model parameter η . The second line of Equation A7 means that the accumulating elements cannot take on negative values. \mathbf{x}_n continues to be updated until one of the activation values exceeds its threshold or until the recall period ends.

Because we do not consider reaction times directly here, we could have chosen a simpler instantiation of the recall process without as many free parameters. We chose to include this more complex decision process so that CMR2 embeds CMR (Polyn et al., 2009) and TCM-A (Sederberg et al., 2008). This also leaves open the possibility to consider reaction times in future CMR2 simulations.

CMR2 incorporates repetition suppression (Brown, Preece, & Hulme, 2000; Farrell & Lewandowsky, 2008; Lewandowsky, 1999; Lewandowsky & Murdock, 1989; Page & Norris, 1998) by assuming that at the beginning of recall, each item i begins with its threshold $\Theta_i = 1$ (the fixed threshold value used for CMR). After i is retrieved, its threshold is set to a maximal value that decreases as a function of the number of subsequent retrievals, j :

$$\Theta_i = 1 + \omega\alpha^j. \tag{A8}$$

ω is a model parameter whereby lower values increase the probability of retrieving an item again. α is a parameter limited to the range $[0,1]$, such that high values of α correspond to longer lags between which a previously retrieved item can be retrieved again.

When an item wins the recall competition, its representation is reactivated in F , and the item's retrieved context is incorporated into the current context representation. Context is updated as during study (Equation 1), although the context drift rate during recall, β_{rec} , may differ from the context drift rate during encoding, β_{enc} .

Next, the retrieved items' context is compared with the currently active context representation according to Equation 2. This is then used to filter out recalls whose context is too dissimilar or similar to the current context representation. In CMR2, it is assumed that each time step lasts 10 ms, and thus CMR2 assumes that the recall period is modeled as a series of competitions, which continues until CMR2 runs out of "time," set to the same time limit as in each simulated experiment.

Between the end of a recall period and the start of the next list, CMR2 simulates the change in study mode by presenting an additional item to the model using Equation 1 and drift rate β_{post}^{recall} . These between-list items do not form associations with context and do not enter the recall competition.

(Appendices continue)

Appendix B

Details of the Parameter-Fitting Technique

To determine the best-fitting parameter set for each simulation, we used a genetic algorithm search that minimized fitness values, quantified as the sum of squared errors between CMR2 predictions and experimental data weighed by the *SE* in the experimental data. For each generation, all parameter sets were ranked based on their fitness values. The top-ranking 20% of parameter sets were used as parents for the next generation, whereby all parameter values were mutated and randomly assigned to new parameter sets.

We determined an initial generation of 2,000 parameter sets by randomly selecting a set of 2,000 values for each parameter, drawn from a uniform distribution of a predetermined range. We then determined four generations of 2,000 parameter sets where the additive mutation noise for each parameter was drawn from a normal distribution of mean 0 and *SD* equal to 20% of each parameter range. Then, we ran 10 generations of 1,000 parameter sets where the mutation rate was 10% of each parameter range. Next, we ran five generations of 500 parameter sets, where the mutation rate was 5% of the parameter range, and each parameter set generated three times as much data as the original experiment. Lastly, we reran the top 250 parameter sets, using 10 times the experimental data. From this final generation, the parameter set with the smallest fitness value was deemed the best-fit parameter set. For each simulation, different aspects of the experimental data were used to assess the fitness value, as listed below.

Simulation 1

The number of PLIs recalled per trial; the proportion of PLIs recalled for list-lags of 1, 2, and 3; the serial position curve; the probability of first recall for the final three serial positions; the lag-conditional response probabilities for lags from -5 to 5 , excluding the first two valid output transitions.

Simulation 2

The number of PLIs recalled per trial; the probabilities of rejection for PLIs and correct recalls; the proportion of nonrejected PLIs recalled for list-lags of 1, 2, and 3; the serial position curve; the probability of first recall for the final three serial positions.

Simulation 3

Proportion of recalled target-list items as a function of target-list length, intervening list-length, and task between lists; proportion of recalled intervening-list intrusions as a function of task between lists.

Simulation 3a: Target-List Recalls With RI

In this parameter search, we sought to find a parameter set such that intervening list-length impacted target-list recalls even with recall between lists (i.e., the top and bottom panels of Figure 5 would each look like the top panel of Figure 5). We fit the same 10 data points as in Simulation 3 above, except we set the proportion of target-list items with recall between lists to be equal to the proportion of target-list items without recall between lists. These four data points (2 intervening list-lengths \times 2 target list-lengths) were matched for intervening list-length and target list-length.

Simulation 4

For $n \in \{1, 2, \dots, 8\}$: the proportion of recall from experimental lists for each $3n-2$; the proportion of recall averaged from experimental lists for each pair of $3n-1$ and $3n$; the proportion of recall from control lists for each $3n-2$; the proportion of recall averaged from control lists for each pair of $3n-1$ and $3n$.

Appendix C

Experimental Methods

The new experimental data reported in the manuscript were collected as part the Penn Electrophysiology of Encoding and Retrieval Study (PEERS), involving three multi-session experiments that were sequentially administered. The methods described here refer to the younger subjects (age 17–30) who took part in Experiment 3 of PEERS. These subjects consisted of students and staff at the University of Pennsylvania, Drexel University, Rowan

University, Temple University, University of the Arts, and the University of the Sciences. All subjects completed Experiment 1 (Lohnas & Kahana, 2013) and Experiment 2 (Lohnas & Kahana, 2014) of the PEERS study before participating in Experiment 3. This experiment consisted of a between-subjects manipulation: immediate free recall (IFR, $n = 51$) or externalized free recall (EFR, $n = 92$).

(Appendices continue)

Immediate Free Recall

Subjects participated in four separate sessions. Each session consisted of 16 lists of 16 words presented one at a time on a computer screen. Each study list was followed by an IFR test and each session ended with a recognition test. Half of the sessions were randomly chosen to include a final free recall test before recognition, in which subjects recalled words from any of the lists from the session. The recognition and final free recall manipulations are not considered here.

Words were either presented concurrently with a task cue, indicating one of two judgments that the subject should make for that word, or with no encoding task, although the manipulation of encoding task was not considered here. The two encoding tasks were a size judgment (“Will this item fit into a shoebox?”) and an animacy judgment (“Does this word refer to something living or not living?”), and the current task was indicated by the color and typeface of the presented item. Using the results of a prior norming study, only words that were clear in meaning and that could be reliably judged in the size and animacy encoding tasks were included in the pool. There were three conditions: no-task lists (subjects did not have to perform judgments with the presented items), single-task lists (all items were presented with the same task), and task-shift lists (both types of judgments were used in a list, although each item was presented with only one judgment type).

Lists were also constructed such that varying degrees of semantic relatedness occurred at both adjacent and distant serial positions, although we collapsed recalls across the semantic conditions. Semantic relatedness was determined using the Word Association Space (WAS) model (Steyvers, Shiffrin, & Nelson, 2004). WAS similarity values were used to group words into four similarity bins (high similarity, $\cos\theta$ between words > 0.7 ; medium high similarity, $0.4 < \cos\theta < 0.7$; medium-low similarity,

$0.14 < \cos\theta < 0.4$; low similarity, $\cos\theta < 0.14$). Two pairs of items from each of the four groups were arranged such that one pair occurred at adjacent serial positions and the other pair was separated by at least two other items.

Words were drawn from a pool of 1,638 words taken from the University of South Florida free association norms (Nelson, McEvoy, & Schreiber, 2004; Steyvers et al., 2004; available at http://memory.psych.upenn.edu/files/wordpools/PEERS_wordpool.zip). Each item was on the screen for 3,000 ms, followed by jittered 800–1,200 ms interstimulus interval. If the word was associated with a task, subjects indicated their response via a keypress. After the last item in the list, there was a 1,200–1,400 ms jittered delay, after which a tone sounded, a row of asterisks appeared, and the subject was given 75 s to attempt to recall any of the just-presented items.

Externalized Free Recall

The EFR methods were identical to the IFR manipulation except for the following. The EFR procedure was introduced in a preliminary session which began identically to the IFR condition. After the third list, instructions appeared on the computer screen indicating that subjects should additionally say aloud every time a specific, salient word came to mind while performing free recall. Subjects were also instructed to press the spacebar immediately after recall of an intrusion or repetition. After this preliminary session, subjects performed five experimental sessions with the same methods except that subjects were given EFR instructions at the beginning of each free recall session.

Received March 14, 2013

Revision received January 22, 2015

Accepted January 27, 2015 ■