Psychonomic Bulletin & Review
2005, 12 (5), 787-805

Going beyond a single list: Modeling the effects of prior experience on episodic free recall

YEVGENIY B. SIROTIN
Columbia University, New York, New York

DANIEL R. KIMBALL
University of Texas, Arlington, Texas

and

MICHAEL J. KAHANA
University of Pennsylvania, Philadelphia, Pennsylvania

We present an extension of the search of associative memory (SAM) model that simulates the effects of both prior semantic knowledge and prior episodic experience on episodic free recall. The model incorporates a memory store for preexisting semantic associations, a contextual drift mechanism, a memory search mechanism that uses both episodic and semantic associations, and a large lexicon including both words from prior lists and unpresented words. These features enabled the model to successfully account for the effects of prior semantic knowledge and prior episodic learning on the pattern of correct recalls and intrusions observed in free recall experiments.

In this article, we present a model that simulates the effects of prior experience on episodic free recall, including the effects of previously learned semantic relations and prior episodic learning. The model presented here, which we term *eSAM*, is an extension of the search of associative memory (SAM) model (Gillund & Shiffrin, 1984; Raaijmakers & Shiffrin, 1981; Shiffrin & Raaijmakers, 1992). SAM is an associative model of memory in which it is posited that, during study, list items become associated with each other and with the study context in proportion to the amount of time the items spend in a limited-capacity rehearsal buffer. SAM further assumes that retrieval is cue dependent, with the list context and previously recalled items serving as retrieval cues for other items, and the probability of retrieving an item being determined by strength-dependent competition among all items associated to a given set of cues. SAM has been applied to a broad range of free recall phenomena, including the effects of presentation rate and list length (Raaijmakers & Shiffrin, 1980), part-set cuing (Raaijmakers & Shiffrin, 1981), word frequency (Gillund & Shiffrin, 1984), interference and forgetting (Mensink & Raaijmakers, 1988), list strength (Shiffrin, Ratcliff, & Clark, 1990), generation (Clark, 1995), and temporal contiguity (Kahana, 1996).

Notwithstanding SAM's far-ranging ability to simulate recall phenomena, instantiations of SAM to date have made certain simplifying assumptions that impair a test of whether the model can simulate the effects of prior experience on free recall. A vast body of empirical evidence suggests that episodic recall of a list learned during an experiment is affected both by preexperimental semantic relations involving list items (see, e.g., Deese, 1959; Glanzer, Koppenaal, & Nelson, 1972; Howard & Kahana, 2002; Kahana & Wingfield, 2000; Pollio, Richards, & Lucas, 1969; Roediger & McDermott, 1995; Romney, Brewer, & Batchelder, 1993; Tulving, 1968) and by the prior study of other lists during the experiment (e.g., Anderson & Bower, 1972; Kahana, Howard, Zaromb, & Wingfield, 2002; Postman & Underwood, 1973; Tulving, 1966; Zaromb et al., 2005).

Although the general theory underlying SAM allows for the storage in memory of associations formed prior to the study of a list, previous applications of the model to recall have represented only the items appearing in, at most, two lists presented experimentally for recall, and they generally have assumed that the only relations involving such list items are those that arise during study and recall of the list.

The eSAM model implements the SAM theory more generally by avoiding many of the simplifying assumptions of earlier work. The eSAM model extends the SAM framework to address the effects of prior experience by incorporating four major features. First, eSAM explicitly represents preexperimental, pairwise semantic associations among words. Although eSAM is a priori neutral as to the best measure of semantic association strength for this purpose, we use two particular measures of seman-

The authors acknowledge support from National Institutes of Health Research Grant MH55687 and Training Grant NS07292. Thanks are due Franklin Zaromb for useful discussions during the early stages of this project and Mark Steyvers and Simon Dennis for providing us with access to the WAS and LSA norms. Correspondence concerning this article should be addressed to M. J. Kahana, Department of Psychology, University of Pennsylvania, Suite 302C, 3401 Walnut St., Philadelphia, PA 19104 (e-mail: kahana@psych.upenn.edu).

tic relatedness, because they have previously been applied to large corpora of words: latent semantic analysis (henceforth, LSA; Landauer & Dumais, 1997) and word association space (henceforth, WAS; Steyvers, Shiffrin, & Nelson, 2005). Second, eSAM searches memory at retrieval, using a combination of these preexperimental semantic associations and the episodic associations formed during study. Third, eSAM incorporates a mechanism by which context changes from list to list, so that the context associated with any given list is a strong cue for retrieving items from that list but is also a weaker retrieval cue for items appearing in earlier lists (cf. Mensink & Raaijmakers, 1988). Fourth, eSAM uses a large lexicon that includes words beyond those on the lists presented during an experiment.

In the next section, we will discuss several experimental results showing the effects of semantic relations and prior episodic learning on free recall of the most recently presented (target) list. Subsequent sections will describe the architecture and operations of the SAM model, then the features modified or added by the eSAM model, and finally, results demonstrating eSAM's ability to simulate the effects of prior experience, as well as certain standard free recall effects.

EFFECTS OF PRIOR EXPERIENCE ON EPISODIC MEMORY

The influence of preexisting semantic relations on episodic recall has been of interest to psychological researchers at least since Bartlett (1932). One phenomenon in this vein is subjects' tendency to use semantic relations to guide item-to-item transitions during recall. A key finding is that, when recalling categorized word lists, subjects tend to recall words from the same natural category together, even when presentation order is randomized (Bousfield, 1953; Glanzer, 1976; Jenkins & Russell, 1952; Kahana & Wingfield, 2000; Pollio et al., 1969; Wingfield, Lindfield, & Kahana, 1998). Similarly, Romney et al. (1993) reported that when subjects recalled a list comprising items from a single category, category members that were more closely related to each other also tended to be recalled more closely together in the output sequence; that is, subjects chose a shorter path through multidimensionally scaled semantic space than would be expected at random. In addition, Howard and Kahana (2002), using study lists consisting of unrelated words, showed that the stronger the semantic relation between two list words, the more likely it was that they would be recalled in neighboring output positions. Howard and Kahana (2002) also reported that the stronger the semantic association between two recalled words in neighboring output positions, the shorter the response latency was between the two words.

In addition to these effects of semantic relations, prior list learning also affects episodic free recall. In a multiple list learning paradigm, subjects sometimes (albeit infrequently) intrude words from lists that have been presented in the experiment previously (prior-list intrusions, PLIs), as well as words not presented on prior lists or the current list (extra-list intrusions, XLIs). PLIs tend to come from the most recently presented prior list, and the probability of a PLI from a particular list falls off rapidly as the list recedes further into the past (Kahana et al., 2002; Murdock, 1974). Zaromb et al. (2005) found that when lists contained mixtures of novel items and items repeated from earlier lists, subjects recalled repeated items more often than novel items, but recalls of repeated items were also more likely to be followed by PLIs than were recalls of novel items.

THE SAM MODEL

In this section, we will describe the simplified simulation recall model first reported in Raaijmakers and Shiffrin (1980, 1981), which forms the foundation of SAM, as well as a number of subsequent modifications to that model, many of which are incorporated into eSAM.

Memory Stores

The SAM model assumes the existence of two memory stores: short-term memory (STM) and long-term memory (LTM). Within STM, rehearsal processes are idealized in the form of a limited-capacity buffer in which studied words become associated through a rehearsal process, as will be described below. LTM contains values for the strengths of the pairwise associations among words, as well as associations between each word and the list context. List context is conceptualized as the temporal and situational setting for a particular list. $S(i, j)$ denotes the strength of association between words i and j, and $S(i, \text{context})$ denotes the strength of association between word i and the list context. For the sake of simplicity, all associations in LTM are assumed to be episodically created in the course of rehearsal during study, so the strengths in LTM are set to zero prior to study (although these associative strengths are later reset to a residual value for pairs of words that are not rehearsed together during study). The general SAM theory allows for preexperimental semantic associations among words, but this feature of the theory has not been implemented in a simulation model to date, and therefore, there has been no test of the theory's capacity to model the effects of such associations on episodic free recall. The only semantic associations that have been modeled thus far have involved categorized lists (Gronlund & Shiffrin, 1986; Raaijmakers, 1979; Raaijmakers & Shiffrin, 1980). However, in these instances, although the associations between category names and category exemplars were represented, there were no representations of the pairwise semantic associations between exemplars in the same or different categories or between a category name and an exemplar classified in a different category. Indeed, it was assumed that an item was either a member of a category or not and that an item could not be a member of more than one category. These simplifying assumptions run counter to evidence that category membership is graded (see, e.g., Z. Estes, 2003). As will be seen, eSAM handles categorized lists quite differently.

Storage

SAM assumes that, during study of a list, each list item enters the STM buffer as it is presented and that a subject rehearses the items occupying the buffer at any given time, thereby increasing the strengths of the items' episodic associations in LTM. In particular, rehearsal increases the strength of association between each item in the buffer and the list context; for each unit of time that an item spends in the buffer, the strength of its association to context is incremented in LTM by the value of parameter a. Rehearsal also increases the strength of the association in LTM between any two items that simultaneously occupy the buffer; for each unit of time that two items spend together in the buffer, their interitem strength is incremented by the value of parameter b.

Kahana (1996) substituted two parameters in lieu of b, one being used to increment interitem strengths in the forward direction—that is, from earlier-presented items to later-presented items (b_1)—and the other being used to increment strengths in the backward direction, from later-presented items to earlier-presented items (b_2). This enabled Kahana to simulate the bias in output order that favors item-to-item recall transitions in the forward direction (further discussed in connection with Simulation 1 below). The eSAM model incorporates these forward- and backward-incrementing parameters.

SAM also represents the association of an item to itself—that is, autoassociation—and includes parameters that increment an item's autoassociative strength when the item occupies the buffer in STM during study (parameter c) and when it is output during recall (parameter g). Autoassociative strength and its incrementing have played a role in the modeling of recognition memory (see Gillund & Shiffrin, 1984), but not of free recall. Accordingly, all autoassociative strengths and parameters c and g are set to zero in our simulations.

The amount of time that each item spends in the STM buffer during study is determined by the presentation rate, the size of the buffer (the maximum number of items that can simultaneously occupy the buffer), and the rule for displacement of items from the buffer. In Raaijmakers and Shiffrin (1980) and other early implementations, the size of the STM rehearsal buffer, r, was set at a single fixed value for all subjects, with $r = 4$ typically providing the best fit to free recall data. Kahana (1996) found it useful to allow the size of the buffer to vary for each subject, with r being randomly selected from a distribution having a mean of μ_r and a standard deviation of σ_r. The eSAM model incorporates this mechanism for varying r.

Once the buffer is full, each new item displaces one of the items then occupying the buffer. The general SAM theory is silent concerning the particular rule governing displacement. The simulation model of Raaijmakers and Shiffrin (1980) assumed that each item in STM had an equal probability of being displaced by the new item. Kahana (1996) found that an alternative displacement rule proposed by Phillips, Shiffrin, and Atkinson (1967) provided a better fit to data on free recall. The Phillips et al. rule assumes a bias in favor of displacing items that have been in the buffer longer than others. Under this rule, the probability that the ith buffer item is to be displaced is given by

$$\frac{q(1-q)^{i-1}}{1-(1-q)^{r}},$$

where q is a fixed parameter of the model that determines the degree of bias favoring displacement of older items. Later-presented items occupy lower ordinal positions in the buffer than do earlier-presented items, thus ensuring a bias under the displacement rule favoring displacement of earlier-presented items. The eSAM model incorporates the Phillips et al. displacement rule.

For a pair of list items that are never rehearsed together in the STM buffer during study, SAM assigns the pair a residual interitem strength of association in LTM, equal to the value of parameter d. As will be discussed later, eSAM incorporates the general concept of residual episodic strength but implements the concept somewhat differently.

Retrieval From STM and Search of LTM

SAM posits a two-stage retrieval process for immediate free recall, the first stage reflecting the output of items in the STM buffer at the beginning of recall, and the second stage reflecting retrieval of items from LTM. According to SAM, items in the STM buffer are always available for immediate recall, so the items in the buffer at the end of list presentation are output first during immediate free recall. In delayed free recall, STM is assumed to have been emptied during the retention interval, and recall therefore begins with retrieval from LTM. The simulation model of Raaijmakers and Shiffrin (1981) assumed that, following the end of the study period, the buffer emptied at the same rate as that for items displaced during study. This mechanism enables SAM to account for data showing that delaying free recall eliminates positive recency but does not generate negative recency (Postman & Phillips, 1965). The eSAM model uses a similar poststudy displacement rule.

Retrieval of items from LTM results from a strength-dependent competition among all items associated to a given set of cues. Each cycle of the search process includes at least two phases: First, an item is *sampled*, and then it may or may not be *recovered*—that is, identified as a particular word. An additional output decision step has sometimes been inserted following recovery. For example, Mensink and Raaijmakers (1988) added a postrecovery decision step to determine whether a recovered item had occurred in the present list or not (see also Raaijmakers, 2003). In eSAM, a similar step is added when items are repeated across lists (see Simulation 4 below).

SAM begins the search of LTM by using context as a retrieval cue. The probability of sampling an item i when using context alone as a retrieval cue is

$$P_s(i \mid \text{context}) = \frac{S(i, \text{context})}{\sum_{k=1}^{N} S(k, \text{context})},$$

where N is the total number of items stored in LTM. This equation ensures that items with greater strengths of association to the list context are more likely to be sampled. Once sampled, the probability that item i is recovered is

$$P_r(i \mid \text{context}) = 1 - e^{-S(i, \text{context})}.$$

Thus, recovery also depends on the strength of association between the item and the list context.

If an item is recalled, it is then used in combination with context to cue recall of another list item. In this case, the probability of sampling item i, given that both context and the just-recalled item j serve as retrieval cues, is

$$P_s(i \mid j, \text{context}) = \frac{S(i, j)S(i, \text{context})}{\sum_{k=1}^{N} S(k, j)S(k, \text{context})},$$

and the probability of recovering the item is

$$P_r(i \mid j, \text{context}) = 1 - e^{-S(i, j) - S(i, \text{context})}.$$

However, regardless of an item's strength of association to the context, to other retrieved items, or to both, the item cannot be recovered if the same retrieval cues failed to recover the item previously or if the item has previously been recovered.

When retrieval cues are successful in recovering an item, the strengths of the item's associations to the retrieval cues are incremented in LTM. The strength of association between the recovered item and the list context is incremented by the value of parameter e, and the strength of association between the recovered item and any other item then being used as a retrieval cue is incremented by the value of parameter f. Thus, different parameters are used at test than at encoding to increment associative strengths in LTM. As with the incrementing of interitem strengths at study, eSAM adopts Kahana's (1996) bifurcation of parameter f into f_1 for forward associations and f_2 for backward associations.

Stopping Rules

There are two rules determining when a subject stops searching, one governing when a subject abandons search with a particular set of retrieval cues, and a second governing when the subject abandons search altogether. When there have been $L_{\max}$ consecutive failures at recovery, using a particular item together with context as retrieval cues, SAM assumes that the subject reverts to using the context alone as a retrieval cue. Search stops altogether when $K_{\max}$ recovery failures have accumulated over all sets of retrieval cues.

Contextual Drift

Mensink and Raaijmakers (1988) added a mechanism to SAM that allowed for the change of context across multiple study and recall episodes. In accord with the classic stimulus sampling theory of W. K. Estes (1955a, 1955b), Mensink and Raaijmakers represented context as a set of elements, with each element in either an active or an inactive state at any given time. The identity of the active contextual elements changed over time—that is, context drifted—with some active elements becoming inactive and some inactive elements becoming active at each time step. At a given time step, associations between active contextual elements and items then in STM were strengthened. Memory was probed using the contextual elements active at the time of test. Therefore, the probability that an item would be sampled and recovered was proportional to the number of contextual elements that were active at both the time of encoding the item and the time of test. By incorporating contextual drift, Mensink and Raaijmakers were able to use SAM to simulate various interference and forgetting effects observed in paired-associate experiments. The eSAM model uses a related mechanism for contextual drift, discussed below.

THE eSAM MODEL: SAM With Preexisting Memory

Semantic and Episodic Associations in LTM

Semantic matrix. The eSAM model incorporates a separate semantic matrix (SM) that stores preexperimental semantic associations between each pair of words in the lexicon. The strength of the associations in the SM remains fixed during the course of the experiment, reflecting an assumption that semantic associations are not significantly affected by episodic experience on the scale of a single experiment (see also Nelson, McKinney, Gee, & Janczura, 1998).

The eSAM model is a priori neutral as to the best measure of semantic relatedness. In each of the simulations that follow, we used two metrics for this purpose: LSA (Landauer & Dumais, 1997) and WAS (Steyvers et al., 2005). Each of these two metrics provides a pairwise measure of semantic relatedness for a large number of words, which is a useful property in a model such as eSAM. However, the two metrics use somewhat different procedures to measure semantic relatedness.

LSA assumes that words that are related in meaning tend to occur close together in texts. The method begins by taking a large corpus of text and counting the number of times that a given word i occurred in a given paragraph j. The resulting matrix, $L(i, j)$, has as many rows as there are words in the corpus and as many columns as there are paragraphs. A mathematical technique called *singular value decomposition* (SVD) is then used to transform the matrix in such a way as to reduce the number of columns while preserving the similarity structure among the rows. Semantic relatedness is measured by the cosine of the angle between vectors consisting of the entries in a particular pair of rows ($\cos\theta$). In our simulations, we used the LSA values computed from texts that a typical person would be likely to have encountered between the third grade and the first year in college. (For a more thorough treatment and discussion, see Landauer & Dumais, 1997; Landauer, Foltz, & Laham, 1998.)

WAS uses SVD to transform the free association word norms collected by Nelson, McEvoy, and Schreiber (2004) into a multidimensional semantic space. The asymmetric associative strengths given by the norms are made symmetric by summing the forward and backward associative strengths (the tendency of *dog* to evoke *cat* and the tendency of *cat* to evoke *dog*). Each word is then represented as a vector of its strengths of associations to other words. SVD is then applied to reduce the dimensionality of the resultant matrix. The relatedness of two words can be calculated as the cosine of the angle between their vectors in semantic space. With this method, words that are directly associated or that share associates have large cosine values. Words that do not directly share associations may still have high relatedness values because of their indirect associations. This method is discussed in greater detail in Steyvers et al. (2005).

Episodic matrix and contextual drift. The other component of LTM is an episodic matrix (EM), which stores the interitem and item-to-context associations formed during study—that is, the associations that made up the whole of LTM in previous versions of SAM. The values for the episodic associative strengths in EM are initialized to random values drawn from a normal distribution with a mean of μ and a standard deviation of σ, which are fixed parameters of the model. The episodic associative strengths in EM are incremented during study and recall in the same way as in previous SAM versions. For items that are not jointly rehearsed in the STM buffer during study, the initialized EM values serve the same function as the residual LTM values in SAM.

In eSAM, contextual drift occurs through probabilistic exponential decay of the item-to-context strengths after a subject finishes with each list, such that

$$S(i,\text{context})_l = \rho S(i,\text{context})_{l-1} + \varepsilon, \quad (1)$$

where l is a counter that represents the index of the current list, $S(i, \text{context})_{l-1}$ and $S(i, \text{context})_l$ are the levels of item-to-context strength for lists $l-1$ and l, respectively, ρ represents the proportion of the item-to-context strength that is conserved between lists, and ε is a noise term taken from a normal distribution with a mean of μ and a standard deviation of σ.

This contextual drift mechanism, which takes on a simple autoregressive form, is simpler than that implemented by Mensink and Raaijmakers (1988), inasmuch as a more complex drift mechanism did not appear necessary to simulate the phenomena considered here. In the General Discussion section, we will consider the relative merits of these two mechanisms.

Use of Semantic, Episodic, and Contextual Associations During Retrieval

When context and the previously recalled item are used as retrieval cues, the sampling and recovery processes in eSAM involve the use of both the semantic associations in SM and the episodic and contextual associations in EM. Inspired by the general retrieval formulae offered by Raaijmakers and Shiffrin (1980), we adopted a sampling rule that multiplicatively combines the strengths of semantic, episodic, and contextual associations to determine probabilities of sampling when LTM is searched using the context and a recalled item as retrieval cues. Accordingly, the probability of sampling item i following the recovery of item j is

$$P_s(i \mid j,\text{context}) = \frac{S_s(i,j)^{W_s} S_e(i,j)^{W_e} S(i,\text{context})^{W_c}}{\sum_{k=1}^{N} S_s(k,j)^{W_s} S_e(k,j)^{W_e} S(k,\text{context})^{W_c}}, \quad (2)$$

and the probability of recovering the item is

$$P_r(i \mid j,\text{context}) = 1 - e^{-W_s S_s(i,j) - W_e S_e(i,j) - W_c S(i,\text{context})}, \quad (3)$$

where $S_s(i, j)$ is the strength of the semantic association between items i and j in SM, $S_e(i, j)$ is the strength of the episodic association between items i and j in EM, $S(i, \text{context})$ is the strength of the episodic association between item i and context in EM, N is the number of words in LTM, and W_s, W_e, and W_c are the parameters for the weighting of retrieval cues consisting of semantic interitem associations, episodic interitem associations, and item-to-context associations, respectively. The recovery probability for an item retrieved using context alone is calculated using the item's mean semantic and episodic relatedness to other items in memory.

A multiplicative sampling rule is not the only formulation possible. For example, in an earlier version of eSAM, we attempted to incorporate semantic influences on recall by summing EM and SM. In this implementation, semantic effects could be observed after a single trial, but with multiple study–test trials, the strengths in EM increased while the strengths in SM remained fixed. As a result, the model counterfactually predicted that the effect of semantic relatedness on episodic recall would become negligible at late stages of learning.

Moreover, we believe that it is in the spirit of SAM to allow all of an item's associations to the retrieval cues to influence retrieval interactively. Use of a multiplicative retrieval rule means that each type of association—semantic, episodic, and contextual—modulates the influence of the other types of associations on sampling probabilities. Thus, given two items that have relatively high contextual strengths and high episodic strengths of association to a third item but one of which has a higher semantic strength of association to the third item than does the other, the model predicts that the item with the higher semantic strength will be more likely to be sampled when the third item is used as a retrieval cue. However, both of the items may have a higher probability of sampling than do other items, because of their high contextual and episodic strengths, and this advantage may increase if episodic and contextual associations are given more weight during retrieval via higher values for the retrieval weight parameters, W_e and W_c.

To eliminate the potential problem that the multiplicative rule would result in a zero probability of sampling a particular item altogether, we ensured that all of the semantic association strengths were greater than zero. Because the measure of relatedness derived from LSA ranged from -0.0925 to 1 and from -0.1428 to 1 for WAS for the sample of words used in our experiments, we transformed these values by adding an offset to ensure that all strengths were greater than zero. This step seemed reasonable inasmuch as any two words can be related semantically in some way when the pair is considered in isolation. We will further discuss the multiplicative rule in the General Discussion section.

Large-Scale Lexicon

The eSAM model incorporates a lexicon that includes many more words than are presented on a single list, or even during an entire experiment. This large-scale lexicon enables eSAM to represent associations in SM between list words and nonlist words, thus enabling the model to simulate semantically induced XLIs. The larger lexicon also enables eSAM to simulate episodically induced PLIs by including words presented on prior lists that can be retrieved via semantic associations, as well as via contextual associations, subject to the effects of contextual drift.

PARAMETERS

Eleven parameters were varied in fitting eSAM to data. Table 1 lists these parameters and gives the best fitting values obtained in Simulations 1–4, as described below. The first five parameters listed in Table 1 were inherited from SAM. The other six free parameters implemented the features added by eSAM. Three of the new free parameters were used in all four simulations: the three parameters used to weight item-to-context associations (W_c), episodic interitem associations (W_e), and semantic interitem associations (W_s) when LTM is searched during retrieval. Another new free parameter was used in all the simulations except Simulation 1: ρ, the proportion of item-to-context strength conserved following a list presentation and recall episode. The other two new free parameters were added to simulate interlist repetition effects in Simulation 4: R,

Table 1
Best-Fitting Parameter Values for Simulations 1–4

	WAS Simulations				LSA Simulations			
Parameter	1	2	3	4	1	2	3	4
a	0.22	0.13	0.43	0.20	0.18	0.11	0.29	0.13
b_1	0.07	0.26	0.22	0.12	0.06	0.13	0.54	0.05
e	0.11	0.59	0.09	0.04	0.13	0.21	0.20	0.04
f_1	0.10	0.04	0.09	0.09	0.08	0.11	0.10	0.04
K_{max}	25	246	43	29	22	115	54	25
ρ	–	0.99	0.04	0.37	–	0.85	0.05	0.42
W_c	0.59	1.25	1.20	0.59	0.58	1.11	1.35	0.61
W_e	3.22	0.30	1.17	1.77	3.16	1.04	1.00	1.60
W_s	0.82	2.00	3.53	2.12	0.82	4.82	1.06	4.00
R	–	–	–	0.53	–	–	–	0.27
m	–	–	–	0.31	–	–	–	0.25
RMSD	.07	.05	.07	.09	.08	.13	.07	.08

Note—In each simulation, eSAM was fit using either the WAS or the LSA semantic relatedness norms. a is the item–context increment during encoding; b_1 is the forward interitem episodic increment during encoding; e is the item–context increment during recall; and f_1 is the forward interitem episodic increment during recall. K_{max} is the maximum cumulative number of recovery failures during recall. ρ is the proportion of item–context strength conserved across lists (Equation 1). W_c, W_e, and W_s are the retrieval weights for item–context, interitem episodic, and interitem semantic strengths, respectively (see Equations 2 and 3). R and m were used only in Simulation 4: R is the boost in item-to-context strength for repeated items recognized during encoding; m is the item–context strength threshold used for list discrimination (see Equation 5). The following parameters were fixed: the ratio of forward to backward strength increments ($b_1/b_2 = f_1/f_2 = 2$), the distribution of rehearsal buffer size across lists ($\mu_r = 4$, $\sigma_r = 1.4$), the factor biasing displacement of older items from the buffer ($q = 0.266$), the mean and standard deviation of the default value for episodic and contextual strength ($\mu = 0.001$, $\sigma = 0.0005$), the maximum number of retrieval failures using a particular set of cues ($L_{max} = 0.1 \times K_{max}$), and the item–context strength threshold for recognition of repeated items during study ($k = 0.0015$) (see Equation 4). RMSD indicates the root-mean squared deviation between observed and predicted values.

the boost in strength increments for repeated items recognized during study, and m, the threshold of item-to-context strength required for recovery of an item during recall. The Table 1 caption also lists the fixed parameters, along with the values at which they were fixed.

SIMULATIONS

The following simulations fit eSAM to data demonstrating several types of free recall phenomena. In Simulation 1, we verified that eSAM retains the ability to fit certain free recall effects that SAM has previously fit: the effects of list length and presentation rate on the serial position curve (Murdock, 1962) and the effects of interitem temporal contiguity at study on interitem transitions during recall (Kahana, 1996), respectively. Simulations 2 and 4 fit eSAM to two phenomena showing the effects of preexperimental semantic associations on episodic recall—specifically, the effects of category membership (Bousfield, 1953) and semantic relatedness (Howard & Kahana, 2002) on interitem recall transitions, respectively. Simulations 3 and 4 fit eSAM to two sets of phenomena showing the effects of prior episodic learning on free recall of new lists: the pattern of intrusions of items from previously studied lists (Howard & Kahana, 1999) and the effects of repeating items across multiple lists (Zaromb et al., 2005), respectively. We will begin with a description of our general methodology.

General Method

For each fit of the model to a set of data, we used a genetic algorithm (M. Mitchell, 1996) to minimize the root mean squared deviation (RMSD) between observed and predicted values. RMSD has the advantage of being measured in the same units as the dependent measures that are being fit and is interpretable as a global measure of the difference between the model's predictions and the observed data, averaged across all the dependent measures. We first calculated RMSD for each of the curves that were fit in a particular simulation and then took the average of those curve-specific RMSD values, so as not to accord undue weight to curves with more data points than others.

At the start of each fit, a population of parameter sets was generated, with the value of each parameter in a set being randomly selected from a predetermined range of values. This starting population was the first generation of parameter sets created by the genetic algorithm. Each parameter set was used to run eSAM for 500–1,000 simulated subjects, to generate statistically stable predictions.

Those parameter sets from the first generation that had the lowest RMSD values were then used to create the next generation of parameter sets through the processes of mutation and recombination. Mutation creates a particular second-generation parameter set by randomly selecting a particular parameter set from among the best-fitting parameter sets in the first generation; then the value for each parameter in the set is randomly either copied or varied within a specified range. Recombination creates a second-generation parameter set by randomly selecting two first-generation parameter sets as *parents* and, for each parameter, randomly selecting either of the parent's values for that parameter as the child's value. In addition, the best-fitting members of the population were retained from one generation to the next. In each of our fits, these processes iterated through successive generations until the average RMSD value for the population reached an asymptote. The best-fitting parameter set was chosen from among the final generation.

Simulation 1: Effects of List Length, Presentation Rate, and Temporal Contiguity

We simultaneously fit several aspects of the data in Murdock (1962), which reported the effects of list length and presentation rate on the serial position curve in immediate free recall. This curve is characterized by a large recency effect, a small primacy effect, and lower recall of items from middle serial positions. Murdock (1962) varied list length and presentation rate by visually presenting lists of 10, 15, and 20 words at a 2-sec rate and lists of 20, 30, and 40 words at a 1-sec rate, with each list followed by 90 sec of oral free recall. The results showed that increasing the list length or the presentation rate impaired recall of prerecency items but did not affect recall of recency items (Figure 1, left). Raaijmakers and Shiffrin (1980) showed that SAM could fit all six conditions with a single set of parameter values.

We used eSAM to simulate these six serial position curves (as well as the lag-conditional response probability [CRP] curve described below) in a single fit, using a lexicon containing 400 words drawn from the Toronto Word Pool (Friendly, Franklin, Hoffman, & Rubin, 1982) and varying the number of rehearsal cycles between items during study to simulate the 1- and 2-sec presentation rates. Our simulations showed that eSAM fit this pattern equally well with either LSA or WAS as the semantic relatedness metric (Table 1). The results with WAS are presented in the right panel of Figure 1. Fitting all six conditions (and the lag-CRP curve described below) with a single set of parameters, eSAM captured the constancy of the recency effect across conditions and also captured the reductions in recall of prerecency items as list length and presentation rate increased.

We also used the Murdock (1962) data to examine transition probabilities in recall as a means of evaluating the model's use of retrieval cues. One can calculate the probability of recalling an item from serial position i + lag immediately following recall of an item from serial position i—that is, the CRP as a function of lag, termed *lag-CRP* by Kahana (1996). Positive values of lag correspond to forward recalls (i.e., transitions to items from serial positions later than that of i); negative values of lag correspond to backward recalls (i.e., transitions to items from serial positions earlier than that of i). Large absolute values of lag correspond to words spaced widely in the list; small absolute values correspond to words spaced closely together in the list.

To illustrate with an example, if the list had contained the subsequence *absence hollow pupil* and a subject re-

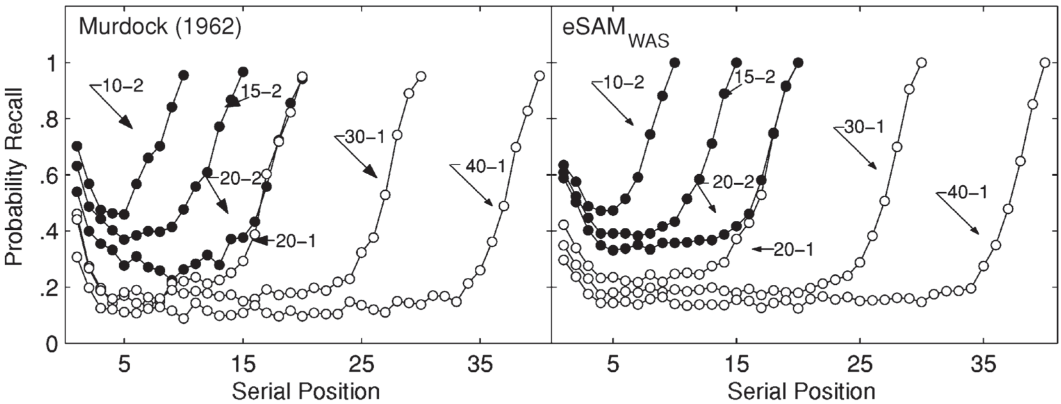

Figure 1. List length and presentation rate effects in immediate free recall. Left panel: data from the six conditions in Murdock (1962). Right panel: simulated results in eSAM using word association space (WAS) as the semantic relatedness metric. For each curve, the first number indicates the list length and the second number indicates the presentation rate (in seconds).

called *hollow* and then *pupil*, the recall of *pupil* would have a lag of +1. If, instead, the subject recalled *hollow* and then *absence*, the recall of *absence* would be associated with a lag of −1. In this case, the subject moved backward in the list. *Absence* followed by *pupil* would yield a lag of +2.

The conditional response probability for a transition of a certain lag is calculated by first tallying the number of times a transition of that lag was made and dividing that tally by the number of times a transition of that lag could have been made. Possible transitions do not include those in which (1) the lag is outside of the bounds of the list or (2) the item has already been recalled. Of course, strictly speaking, these transitions are not impossible, since subjects occasionally make intrusions from outside of the list and sometimes repeat items. However, such transitions are extremely rare in comparison with within-list lags to items not previously recalled. After calculating the lag-CRP function for each subject, these functions are averaged across subjects, and the confidence intervals around the estimated means are calculated.

Kahana (1996) calculated lag-CRP functions for a wide range of experimental conditions. He found that the lag-CRP function has two invariant characteristics. (1) The function decreases monotonically as absolute lag increases, approaching an asymptotic value at large lags; the asymptotic value depends largely on list length, with lower asymptotes for longer lists. (2) For small absolute lags, the function is consistently asymmetric, with an approximately 2:1 ratio of forward to backward recall transitions. The basic pattern has been confirmed in a number of subsequent studies (e.g., Howard & Kahana, 1999; Kahana & Howard, 2005; Kahana et al., 2002; Ward, Woodward, Stevens, & Stinson, 2003; Zaromb et al., 2005). As one example of this pattern, the left panel of Figure 2 shows the lag-CRP function for the 40-item list in Murdock (1962).

In the same simulation in which we fit the serial position curves for the Murdock (1962) data (see Figure 1), we used eSAM to fit the lag-CRP curve for the 40-item list, averaging the separate RMSDs for the lag-CRP and the serial position curves. This simultaneous fitting of the lag-CRP and serial position curves provided a more rigorous test of the model's mechanisms for producing interitem associations than would fitting the curves separately. As in Kahana (1996), we computed the lag-CRP by excluding the first three output positions for both the data and the model. This was done because, in immediate free recall, adjacency and asymmetry are enhanced for the first two to three output positions. This reflects "bleed in" from the recency effect, where the last few list items tend to be recalled as a cluster prior to recall of other items. The lag-CRP is invariant across later output positions, thus reflecting the general tendency to make associative transitions in retrieval from LTM.

By analogy to the recency effect, which illustrates how items near in time to the end of a list are better remembered, Howard and Kahana (1999) referred to associative effects in free recall as illustrating a *lag-recency effect*, since they reveal a preference for recalling items presented near in time to the just recalled item. Very similar lag-recency effects are observed in serial recall (Kahana & Caplan, 2002; Klein, Addis, & Kahana, 2005; Raskin & Cook, 1937).

The eSAM model produced a lag-recency effect that was quite similar to the experimental data, as shown in the right panel of Figure 2. As in Kahana (1996), the asymmetry in the lag-recency effect results from the use of separate parameters for incrementing forward versus backward associations during both encoding and retrieval, as was noted previously. The nonlinear decline in the lag-CRP with increasing absolute lag occurred because the retrieval weighting parameters were allowed to take on values other than 1 (see Equations 2 and 3).

Thus, the modifications to SAM that are incorporated into eSAM do not affect the model's ability to fit the serial position and lag-CRP curves for the Murdock (1962) data, regardless of whether we used WAS (RMSD = .07) or LSA

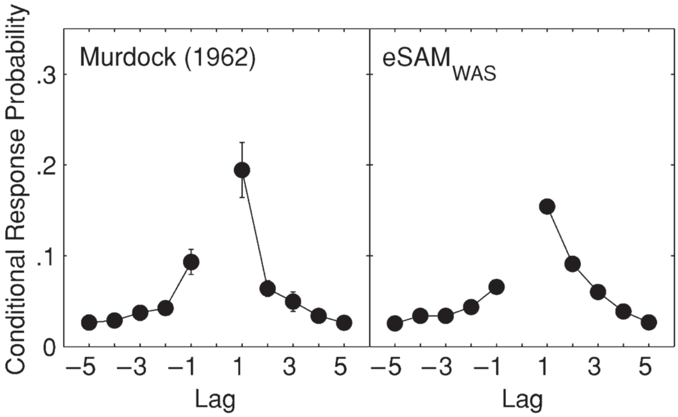

Figure 2. The lag-recency effect in free recall. Left panel: conditional response probability as a function of lag for data from Murdock's (1962) 40-item list condition. Right panel: simulated results in eSAM using word association space (WAS) as the semantic relatedness metric. Error bars denote 95% confidence bands.

(RMSD = .08) as the semantic relatedness metric. We next will evaluate the model's ability to simulate the effects of prior semantic and episodic experience on free recall.

Simulation 2: Category Clustering

To simulate category clustering, we fit eSAM to data from Kahana & Wingfield (2000), in which each subject studied two lists of 20 words, each consisting of four exemplars drawn from each of five natural categories. The subjects studied and recalled a list multiple times until achieving perfect recall, with presentation order randomized anew for each trial. Each subject learned one list consisting of exemplars that were highly prototypical of their categories (high-prototypicality condition) and one list consisting of exemplars that were weakly to moderately prototypical of their categories (low-prototypicality condition). The order of the two lists and the categories from which exemplars were selected were counterbalanced across subjects. Although Kahana and Wingfield tested both young and older subject groups, we fit eSAM only to the younger adults' data, and we used a lexicon of 80 words drawn from those used in Kahana and Wingfield.

As a measure of category clustering, we used the ratio of repetition (Bousfield, 1953), calculated in the following manner: For each recall trial, we calculated the observed ratio of repetition, $RR_{\text{observed}} = r/(n - 1)$, where r is the observed number of intracategory recall transitions (i.e., repetitions) and n is the number of items recalled; we ignored any recalls of an item after its first recall on a given trial. This ratio thus compares the number of intracategory recall transitions to the total number of recall transitions on that trial. However, to obtain a truer measure of the use of category membership to guide recall transitions, we subtracted from the observed ratio an amount equal to the baseline ratio of intracategory transitions that could be expected by chance: $RR_{\text{expected}} = (e - 1)/(ce - 1)$, where c is the number of categories and e is the number of exemplars in each category. Frender and Doubilet (1974) noted that a key advantage of the net ratio, $RR = RR_{\text{observed}} - RR_{\text{expected}}$, as compared with other measures of clustering, is its independence from the number of items recalled on a given trial, thus permitting comparisons across subjects and trials without regard to differences in recall rates that could be affected by subject variables and learning.

Figure 3 (left panel) plots the data for the first three study–recall trials in Kahana and Wingfield (2000), with separate curves for high- and low-prototypicality items. Recall is plotted on the horizontal axis, and ratio of repetition is plotted on the vertical axis. Ellipses represent the 95% confidence region based on a bivariate normal distribution. The figure shows that both recall and semantic clustering, as measured by the ratio of repetition, increased across Trials 1, 2, and 3 for both high- and low-prototypicality word lists, and on each trial, high-prototypicality lists yielded both higher levels of recall and higher levels of clustering than did low-prototypicality lists.

The center and right panels in Figure 3 depict the recall and clustering results from separate eSAM simulations of the Kahana and Wingfield (2000) data, using WAS and LSA as the semantic relatedness metric, respectively. Although both variants of eSAM captured the overall correlation between recall and clustering, WAS provided a much better quantitative fit (RMSD = .05) than did LSA (RMSD = .13). To explain this difference in the fits using WAS and LSA, we examined the treatment of categories in these metrics more closely.

For eSAM to model category clustering on the basis of simple pairwise interitem associations, with no explicit representation of category membership, the semantic metric used must assign higher values for intracategory associations than for intercategory associations. Figure 4 (left panel) shows that, for LSA, the distributions of semantic relatedness values for both intracategory and intercategory pairs are widely dispersed and that they overlap

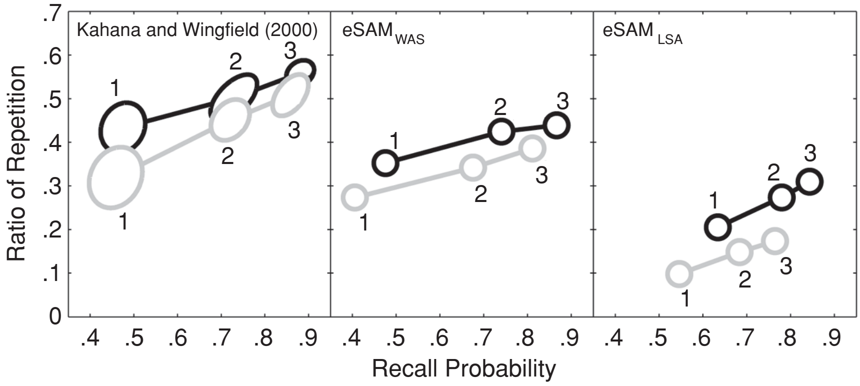

Figure 3. Increases in category clustering with list learning for low-prototypicality lists (gray lines) and high-prototypicality lists (black lines). Ratio of repetition is plotted as a function of recall probability for the first three trials of a multitrial free recall experiment. Left panel: data from Kahana and Wingfield (2000; young subjects). Ellipses represent the 95% confidence region based on a bivariate normal distribution. Middle and right panels: eSAM exhibits similar patterns, although word association space (WAS; middle) better predicts the level of clustering associated with a given level of recall than does latent semantic analysis (LSA; right).

considerably. By contrast, for WAS, the distributions are more highly skewed in opposite directions, with the intercategory pairs bunching close to zero and the intracategory pairs bunching close to 1. As a consequence, and not surprisingly, the receiver-operating characteristic functions for WAS and LSA (Figure 4, right panel) reveal that the discriminability (d') between the intercategory and the intracategory distributions is greater for WAS than for LSA. Thus, the superiority of the WAS fit of the category clustering data seems to be attributable to its superior discrimination between the strengths of intracategory and intercategory associations.

Two mechanisms are likely responsible for the ability of eSAM to cluster responses categorically. First, pairs would be recalled together, due to semantic relatedness. Second, following the initial recall, pairs would become increasingly likely to be recalled together on subsequent trials, due to a strengthening of episodic associations through output encoding.

One final aspect of the simulation to note is the high value for K_{max}, indicating that the model needs many more sampling/recall attempts to fit these data. There is some indication in the literature that, when recalling categorized lists, subjects continue to produce correct responses

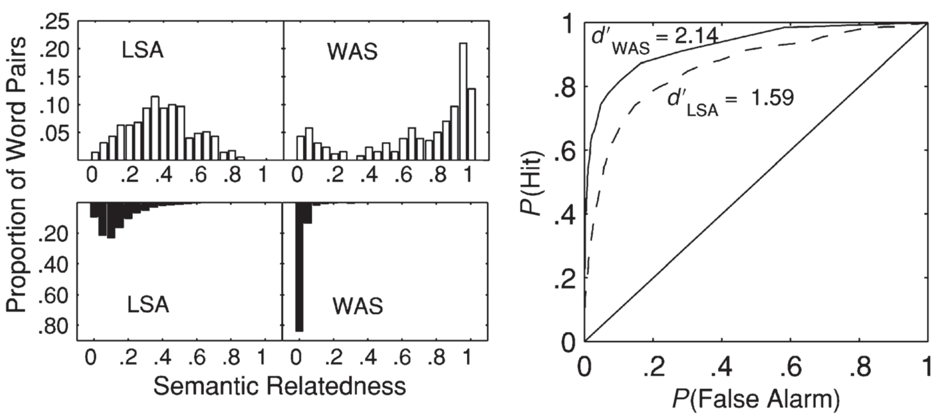

Figure 4. Distribution of within-category and between-category interword relatedness derived from latent semantic analysis (LSA) versus word association space (WAS). The left panels show distributions of within-category (white) and between-category (black) semantic relatedness values for LSA and WAS. The right panel shows receiver-operating characteristic (ROC) functions for the discrimination of within-category and between-category word pairs based on relatedness measures derived using WAS (solid curve) or LSA (dashed curve). These ROC functions show the relation between the hit rate and false alarm rate for classifying word *i* as coming from the same category as word *j* on the basis of the semantic relatedness of words *i* and *j*. If the semantic relatedness (as measured by LSA or WAS) exceeds a threshold, the words are deemed to come from the same category. The associated d' values are reported in the figures.

late into the recall period, especially when the category cue is potent (Roediger, Payne, Gillespie, & Lean, 1982; Wingfield & Kahana, 2002).

Simulation 3: Prior-List and Extra-List Intrusions

In a discussion of some unpublished observations regarding free recall, Murdock (1974) noted that PLIs tend to come from the most recent list and that there is a monotonically decreasing intrusion gradient across earlier lists. In a secondary analysis of data from several free recall studies, Zaromb et al. (2005) showed that the proportional frequency of PLIs decreases sharply with the number of intervening lists since the list on which the intruded item had appeared.

The present simulation fit eSAM to data from Experiment 2 in Howard and Kahana (1999). In that experiment, each subject experienced study-delayed recall trials for 15 different 12-item lists in each of 10 sessions. For some lists, the subjects performed a distractor task for varying periods between item presentations; we did not simulate those lists but only lists with no such interitem distractors. For our analyses of both the behavioral data and the simulation, we excluded recalls of the first 5 lists, so that for each list included in the analysis, the subjects could have intruded items from 5 lists back. We were also interested in the model's ability to simulate XLIs. Because prior instantiations of SAM had not simulated free recall with a lexicon that included words that had not been presented during the experiment, eSAM is the first SAM-type model potentially capable of explicitly modeling the occurrence of XLIs. Accordingly, we simultaneously fit the PLI-recency curve, the average number of PLIs, and the average number of XLIs, as well as the serial position and the lag-CRP curves, as in Simulation 1. The lexicon contained 482 words drawn from the Toronto Word Pool.

As Figure 5 (left panel) shows, PLIs in the experiment tended to come from more recently presented lists than from more remote lists—a PLI-recency effect, consistent with Murdock's (1974) observation. As Figure 5 (right panel) shows, eSAM simulates this pattern reasonably well. Furthermore, Table 2 shows that with the same parameter values, eSAM can reproduce correct means not just for PLIs, but also for correct recalls and XLIs. Using the same parameter values, eSAM also provided a good fit to the serial position and lag-CRP curves (not shown, but reflected in the RMSD value for the fit).

The ability of eSAM both to simulate PLIs and to capture the PLI-recency effect arises primarily due to the operation of the contextual drift mechanism and a relatively greater role accorded to item-to-context associations during encoding and retrieval. The fit generated relatively large values for the *a* parameter (the study phase boost in an item's strength of association to context) and for the W_c parameter (the weight accorded to item-to-context associations in the retrieval process). The fit also generated a relatively small, but nonzero, value for ρ, the parameter signifying the proportion of item-to-context strength that is conserved between lists. The greater role for contextual associations during encoding and retrieval and the nonzero value for ρ enabled the model to simulate PLIs in the first place, and the small value of ρ also ensured that the number of PLIs would be small and that the bulk of the PLIs would come from the most recent list.

The XLIs are made possible by the expanded lexicon and the use of semantic associations to search LTM during retrieval, in combination with residual contextual and episodic associations. These features of the model afford an unpresented item some probability of sampling and recovery, to the extent that it is semantically related to a just-recovered item that is then being used as a retrieval cue.

The low RMSD values obtained in these fits (.07 for both WAS- and LSA-based simulations) attest to eSAM's ability to simultaneously account for many diverse aspects of the experimental data. Here, we are fitting not only serial position curves with great accuracy, but also the within-list associative effects described by the lag-CRP functions, the relative probabilities of correct re-

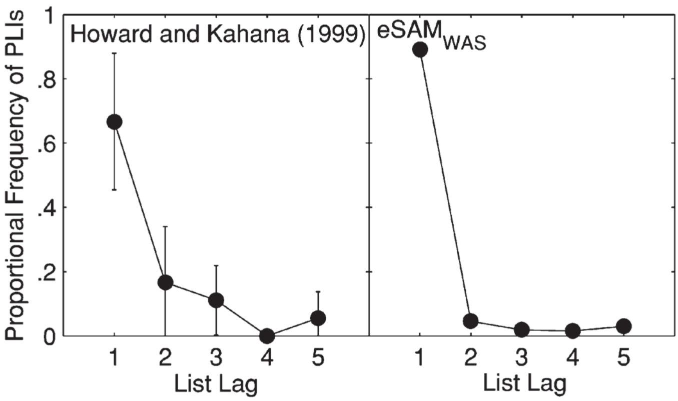

Figure 5. The effect of recency on prior-list intrusions (PLIs). Data from Experiment 2 in Howard and Kahana (1999; left) show that PLIs tend to come from more recent lists. The eSAM model fit to the data (right) captures the quality of this trend, using word association space (WAS). Error bars denote 95% confidence intervals.

Table 2
Average Number of Prior List Intrusions (PLIs), Extra-List Intrusions (XLIs), and Correct Recalls Made for Howard and Kahana (1999) Data and the eSAM Fit Using Word Association Space (WAS)

	Correct	PLI	XLI
Howard and Kahana (1999)	5.0 (0.50)	0.50 (0.2)	0.20 (0.1)
eSAM–WAS	4.7	0.53	0.18

Note—Numbers in parentheses are 95% confidence intervals. Similar results were obtained using latent semantic analysis instead of WAS as the semantic relatedness metric.

calls, PLIs, XLIs, and finally, the temporal gradient of PLIs from previous lists.

Simulation 4: Effects of Semantic Relations and Interlist Repetitions on Recall Transitions

In this simulation, we evaluated eSAM's ability to simultaneously fit several novel features of retrieval in free recall. Focusing on the transitions made during recall, we considered whether eSAM can account for two primary effects of prior experience: (1) the tendency for subjects to make transitions to semantically related items, even in random word lists and (2) the increased tendency to produce PLIs following recall of an item that had appeared both on the current list and on an earlier list.

We fit eSAM to data from Experiment 1 in Zaromb et al. (2005), in which subjects performed 16 study–recall trials on 20-item lists consisting of a combination of new and repeated items. Our primary goals were to fit two aspects of the data: (1) the curve specifying the relative probability of successively recalling items having a particular degree of semantic relatedness (the semantic-CRP curve, described below) and (2) the probability of recalling an item having a particular status—new, repeated, or PLI—as a function of the status of the just-recalled item. In addition, we simultaneously fit the serial position curve, the lag-CRP curve, and the curve specifying the probability of recalling an item first for a particular list as a function of its serial position. The lexicon contained 271 words drawn from the Toronto Word Pool.

Semantic-CRP. Just as the lag-CRP measures the conditional probability of a recall transition as a function of an item's episodic contiguity to the just-recalled item during study, it is possible to measure the conditional probability of a recall transition as a function of an item's semantic similarity to the just-recalled item, as measured using LSA or WAS. Using only the LSA semantic metric, Howard and Kahana (2002) termed this the *LSA-CRP* function. We term it the *semantic-CRP* to reflect the generality of this effect across semantic spaces. This positive relationship is evident in Figure 6 (left panels), which shows the semantic-CRP functions using LSA and WAS for the combined data drawn from the delayed free recall conditions in Howard and Kahana (1999), as well as Kahana et al. (2002) and Zaromb et al. (2005), all three of which experiments exhibited a similar pattern.

Because the slopes of the semantic-CRP functions were nearly identical across the three studies, we simulated the data from Zaromb et al. (2005) alone. eSAM provided a good fit to the semantic-CRP slopes from Zaromb et al., using the same set of parameter values that also simulta-

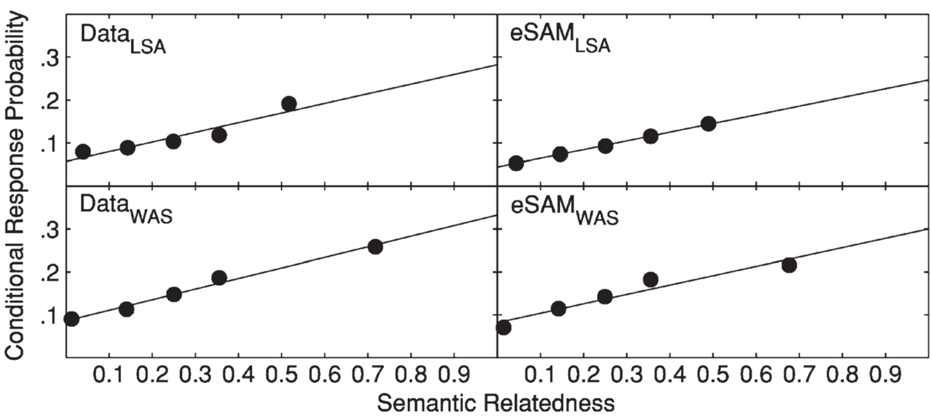

Figure 6. Latent semantic analysis (LSA) and word association space (WAS) semantic relatedness values predict output order in free recall. The distribution of pairwise semantic relatedness was divided into bins containing an equal number of pairs (200 bins for LSA, 1,000 bins for WAS). For each bin, we calculated the probability of a recall transition between two words having a similarity value that fell in the bin, as well as the mean similarity value for such transitions. The bins were further collapsed into five relatedness ranges (less than 0.1, 0.1–0.2, 0.2–0.3, 0.3–0.4, and 0.4–1), with the mean conditional probability and similarity value represented for each range. The conditional response probability is plotted as a function of mean semantic relatedness for LSA (top) and WAS (bottom). The line in each figure represents the fit of a regression applied to the average subject data. The experimental data, taken from Howard and Kahana (1999), Kahana, Howard, Zaromb, and Wingfield (2002), and Zaromb et al. (2005), are from delayed free recall. The regression line slope for the data is .23 (95% CI = 0.02) for LSA and 0.2462 (95% CI = 0.0007) for WAS. For the model, the regression slope is .20 (95% CI = 0.01) for LSA and .218 (95% CI = .009) for WAS.

neously fit the other aspects of the data. This result clearly depends on the use of pairwise semantic associations to guide LTM search during recall. Accordingly, previous instantiations of SAM, which did not include a semantic search mechanism, would not have been able to fit the semantic-CRP.

Interlist repetition effects. The effects of prior episodic learning may also be seen when items are presented on multiple lists within an experimental session. Zaromb et al. (2005) reported both positive and negative effects of such interlist repetitions on recall. Recall of repeated items was enhanced relative to recall of new items. In addition, the enhancement was greater for items that had previously appeared on more recent lists than for those on more remote lists—a list recency effect (see Figure 7, left panel). However, subjects were also more likely to make PLIs following recall of a repeated item than following recall of a new item (see Table 3). Nevertheless, repeated and new items had similar lag-CRP functions, showing similar probabilities of recall transitions to other items from the current list that have particular lags relative to the just-recalled item (see Figure 8, left panel).

To simulate these three interlist repetition effects, we added two features to eSAM: a mechanism that boosts the incrementing of episodic associative strengths for a repeated item while it is in the STM rehearsal buffer during study and a retrieval threshold for item-to-context strength to determine whether an item had appeared in the current target list rather than elsewhere. We next will describe these two mechanisms in more detail, and we then will describe eSAM's fit to data from Zaromb et al. (2005).

Increased episodic strengthening of repeated items during study. The greater overall probability of recalling a repeated item and the list recency effect (see Figure 7, left panel) suggested to Zaromb et al. that subjects may be recognizing the item during study as having appeared in a prior list and according it more attention. An alternative explanation, however, is that the improved recall of items repeated in recent lists is simply a manifestation of recency operating across lists: Just as PLIs exhibit a recency effect (see Figure 5), so too, retrieval of these items will increase the likelihood of recalling a repeated item that was presented on a recent prior list. Zaromb et al. evaluated this account by estimating the increase in recall for repeated items that would be predicted on the basis of the PLI data and found that recency could account for no more than 5% of the increase in performance for repeated items. They therefore concluded that subjects allocate additional attention or encoding resources to items recognized as repetitions. We implement this notion in eSAM by assuming that there is greater strengthening of the repeated item's associations to the study context and episodic associations to other items jointly occupying the STM rehearsal buffer during study.

Accordingly, eSAM assumes that subjects recognize the item as a repeated item if its item-to-context strength exceeds a threshold, such that

$$1 - e^{-S(i,\text{context})} > k, \tag{4}$$

where k is a fixed parameter of the model. If Equation 4 holds true for an item, it is recognized as a repeated item during study, and the normal increments in episodic associative strength are increased by an amount equal to R times the normal increment, where R is a free parameter of the model (see Table 1). Thus, a repeated item's item-to-context strength is incremented by $(1 + R)a$, rather than by a, and its forward and backward interitem strengths are incremented by $(1 + R)b_1$ and $(1 + R)b_2$, rather than by b_1 and b_2, respectively. This recognition mechanism is similar to that proposed by Raaijmakers (2003), who, in simulating spacing effects, also used contextual associations to discriminate previously presented items and also provided for the possibility that the strength of an item would be boosted if it were recognized as having been presented earlier (although no such boost was actually used in the simulations reported by Raaijmakers, 2003).

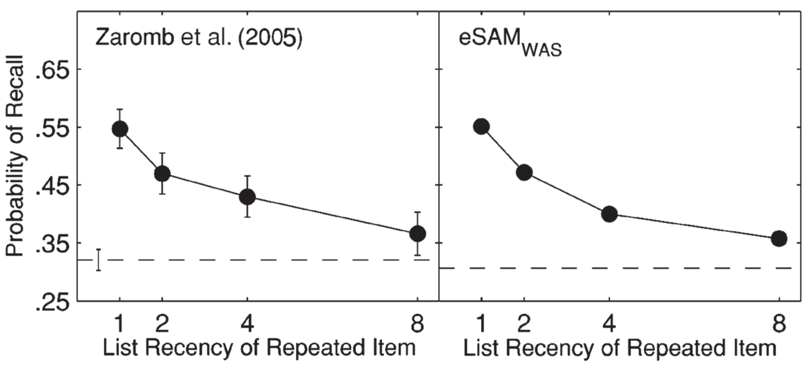

Figure 7. Recall gains resulting from interlist repetition. Left panel: recall probability of repeated items as a function of list recency for data from Zaromb et al. (2005). Right panel: simulated results in eSAM using word association space (WAS) as the semantic relatedness metric. Dashed lines indicate the probability of recalling new items. Error bars denote 95% confidence intervals.

Table 3
Conditional Transition Probabilities Showing the Probability of the Type of Item to Be Recalled Next Given the Type of Item That Has Just Been Recalled

From:	To: New	Repeated	PLI	Stop
Zaromb et al. (2005)				
New	.59 (.02)	.18 (.01)	.09 (.01)	.15 (.01)
Repeated	.61 (.03)	.16 (.02)	.12 (.02)	.12 (.02)
PLI	.33 (.07)	.22 (.08)	.18 (.05)	.28 (.08)
eSAM–WAS				
New	.60	.19	.05	.15
Repeated	.62	.17	.12	.13
PLI	.31	.33	.22	.37

Note—Upper part gives data from Zaromb et al. (2005). Numbers in parentheses are 95% confidence intervals. Lower part presents simulated results in eSAM, using word association space (WAS) as the semantic relatedness metric.

List discrimination during retrieval of lists with repeated items. When items are repeated across lists, subjects may adopt a criterion during recall that permits better discrimination between items that had appeared on previous lists and those that had appeared on the current list. We implemented this notion in eSAM by assuming that subjects adopt an additional criterion for recovery, such that, notwithstanding satisfaction of other recovery criteria, an item is recovered only if its contextual strength exceeds a criterion indicating its presence on the current list—specifically,

$$1 - e^{-S(i,\text{context})} > m, \tag{5}$$

where m is a free parameter of the model (see Table 1).

List recency effects. As is shown in Figure 7 (left panel), Zaromb et al. (2005) reported that the boost in recall probability for a repeated item declined exponentially as more lists intervened between the list on which it originally appeared and the list on which it was repeated, until, after eight intervening lists, the recall probability approximated that of a new item. This pattern is consistent with research on spacing effects in that, although spaced repetition helps recall more than does massed repetition when there is a long retention interval following the second presentation, the reverse is true when the retention interval following the second presentation is short (Glenberg, 1977; Melton, 1963; Peterson, 1966; Peterson, Wampler, Kirkpatrick, & Saltzman, 1963). In Zaromb et al, recall always shortly followed the second presentation, and accordingly, there was better recall of more recently repeated items. Figure 7 (right panel) shows that eSAM is able to fit this pattern, due to its incorporation of a mechanism by which the strength of an item's association to context decays exponentially across succeeding lists.

Recall transitions. Table 3 (top) shows the probabilities of recall transitions to a new item, repeated item, or PLI, conditionalized on the just-recalled item having been a new item, a repeated item, or a PLI. These conditional probabilities were first determined separately for each subject and then averaged across subjects. In calculating each subject's transition probabilities, we collapsed across output position (i.e., first to second response, second to third response, etc.), including only those output positions for which a subject recalled at least two items of a given type. Table 3 (bottom) shows that the model captures this pattern of recall transitions well.

Of particular interest for the eSAM model, both the behavioral and the simulation data in Table 3 indicate that the probability of a PLI is greater following recall of a repeated item than following recall of a new item and greater still following another PLI. Within the eSAM framework, the explanation for this pattern of transitions to PLIs rests in the differing degrees of episodic association strength between these different types of items and PLIs. When the just-recalled item serving as a retrieval cue had been presented both in the current list and in a prior list, episodic associations between that cue item and current list items must compete with the cue item's episodic associations to prior list items. Sometimes the current list items will win

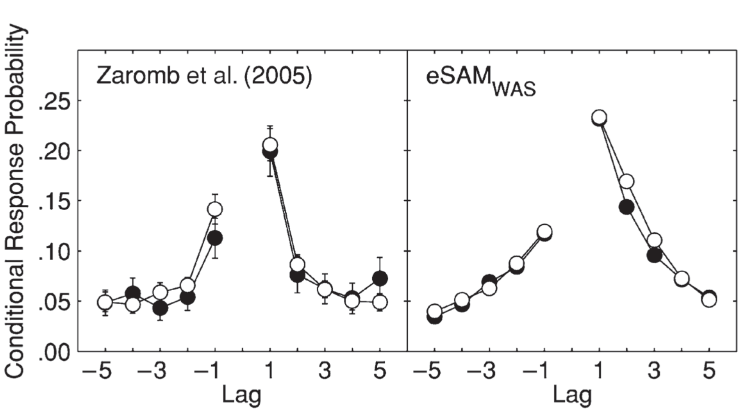

Figure 8. Lag-recency effects for transitions from new and repeated items. Left panel: conditional response probability as a function of lag for new items (filled circles) and repeated items (open circles) in Zaromb et al. (2005). Error bars denote 95% confidence intervals. Right panel: simulated results in eSAM using word association space (WAS) as the semantic relatedness metric.

this competition, and sometimes the PLIs will win, resulting in a moderate level of transitions to PLIs. However, when the just-recalled item is itself a PLI, the retrieval competition is especially likely to be won by a prior-list item, because such an item has substantial episodic associations only to other prior-list items. By contrast, a new item serving as a cue has substantial episodic associations to current list items, but not to unrepeated items from previous lists, and the rate of PLIs will be low in that case, although it will still be nonzero because of transitions based on semantic and residual contextual associations.

Notwithstanding differences in the capacities of new and repeated items to serve as retrieval cues for PLIs and differences in the unconditional probability of recalling a repeated item versus a new item, Figure 8 (left panel) shows that Zaromb et al. (2005) found no distinction between the lag-CRP curves for new and repeated items. The similarity of these patterns suggests that the recall of a repeated item does not significantly affect subjects' tendency to recall nearby list items in succession (the lag-recency effect). Figure 8 (right panel) shows that eSAM captures the similarity in the new and the repeated items' lag-CRP functions.

Comparison of Parameter Values Across Simulations

We did not systematically examine the contributions of the free parameters to each fit, either alone or in combination. However, we can offer some general observations regarding the variation of parameter values across the simulations. For this purpose, Simulation 1, in which we modeled the serial position and lag-CRP curves for the data in Murdock (1962), provides a baseline condition, in that the effects did not depend on semantic relations or on recall of multiple lists. Consistent with the nonsemantic nature of the effects, the free parameter values for Simulation 1, using the WAS and LSA semantic metrics, were quite similar.

Of the remaining three simulations that involved semantic effects and/or recall of multiple lists, Simulation 4 yielded parameter values closest to those of Simulation 1. That simulation involved the presentation of multiple lists, with some items repeated across lists, and several effects were fit simultaneously, including the semantic-CRP curve and several measures that were more episodic in nature (serial position, lag-CRP, probability of first recall curves, and transitions among new items, repeated items, and PLIs). The most notable exception to the similarity between parameter values for Simulations 1 and 4 is the reversal of the values for the episodic and semantic retrieval weight parameters, W_e and W_s, respectively. This reversal makes sense in light of the addition of semantic effects in Simulation 4. As with Simulation 1, the patterns of parameter values for Simulation 4 were quite similar for the WAS and the LSA metrics.

Relative to the parameter values in Simulations 1 and 4, the parameter values to fit the category clustering data in Simulation 2 changed in similar directions for the WAS and the LSA metrics. However, the changes were more dramatic and the quantitative fit was substantially better for the WAS metric than for the LSA metric. The changes reflected the need to achieve the twin goals of ensuring more semantically than episodically based recall transitions and of generating extremely high levels of recall. It seems plausible that, to generate a greater proportion of intracategory than intercategory transitions during recall, the semantic retrieval weight parameter (W_s) remained relatively high, as for Simulation 4, but the episodic retrieval weight parameter (W_e) declined. It also seems plausible that, to achieve an extremely high level of recall, the model compensated for the reduction in the episodic retrieval weight parameter by increasing the values of three contextual association parameters (e, ρ , and W_e) and by prolonging the search of LTM (K_{max} increased).

Finally, Simulation 3 involved modeling multiple list recall and, specifically, the PLI recency effect, the number of PLIs, XLIs, and correct items recalled, and the serial position and lag-CRP curves. Once again, the WAS and the LSA metrics yielded similar patterns of parameter values (with one notable exception, discussed below). As in Simulation 2, the simulated effects appear to have generated competing influences on parameter values. Three of the contextual association parameters were relatively high—a, W_e, and (for the LSA simulation) e—but the fourth, ρ, was very low, albeit nonzero. A plausible interpretation of this pattern is that the first three parameters needed to be high in order to filter out and avoid an excessive number of XLIs, whereas the low value of ρ served to preserve just enough of the high level of contextual association strength across lists to enable eSAM to simulate the proper number and recency of the PLIs. Filtering out XLIs is also likely to be behind the relatively high level of episodic associative strength parameters, b_1 and f_1. The most significant difference in parameter values for the WAS and the LSA simulations was the value of the semantic retrieval weight parameter, which was relatively high for the WAS simulation but relatively low for the LSA simulation. This difference might arise from the difference in the distribution of associations across the two semantic spaces, as discussed earlier, with WAS having a more polarized distribution of association values and LSA having a more uniform distribution. LSA could be more sensitive to semantically induced XLIs due to this distributional property—on average, items tend to be more semantically related to other items—thus requiring a lower semantic retrieval weight to avoid an excessive number of XLIs.

GENERAL DISCUSSION

We have shown that eSAM, an extension of SAM, is able to simulate several key effects of prior semantic and episodic learning on episodic free recall. By including representations of interitem semantic associations in LTM, and by using those associations along with contextual and episodic associations to guide memory search during retrieval, eSAM was able to simulate category clustering, as well as semantically based recall transitions and extra-list intrusions. By also including representations of prior-list items in LTM and adding a decay-based contextual drift mechanism, eSAM was able to simulate correct numbers of PLIs, as well as the PLI-recency effect. Finally, by adding a mechanism to boost the episodic strengthening of prior-list items when they are re-presented on a later list, as well as a mechanism to distinguish between prior-list items and current-list items during recall, eSAM was able to simulate the pattern of effects of such interlist item repetitions on free recall.

Semantic Memory

In eSAM, semantic associations are not affected by events that occur during the experiment, and such associations play a role only during retrieval, acting as an additional source of information that constrains, and is constrained by, contextual and episodic associations in guiding LTM search. There are several reasons for our handling of semantic associations in this way. Perhaps most important, we wanted to maintain the spirit of SAM by extending the sampling rule offered by Raaijmakers and Shiffrin (1980), multiplying semantic association strengths with the contextual and episodic association strengths that are combined multiplicatively in their sampling rule.

In addition, such a multiplicative sampling rule seems useful because it provides a mechanism by which episodic/contextual factors constrain semantic search, thereby avoiding an explosion of semantically induced intrusions, and semantic factors also constrain episodic search, so as to account for effects based on semantic properties of stimuli and/or semantic strategies.

Although the multiplicative rule requires that each type of association always play some modulating role in the sampling process, the relative importance of these different types of associations is adjusted in eSAM through variation of the retrieval weight parameters, which are exponents of the associative strengths in the sampling rule. For example, in Simulation 1, the value of the episodic retrieval weight parameter W_e was much greater than that of the semantic retrieval weight parameter W_s, befitting an episodically oriented task using stimuli that had relatively low interitem semantic associations. By contrast, in the other three simulations, which involved the modeling of categorical clustering and semantically induced intrusions and recall transitions, W_s was greater than W_e. In this way, eSAM is able to emphasize semantic or episodic factors in its search of LTM.

Another aspect of the treatment of semantic associations in eSAM is the use of separate episodic and semantic matrices. This feature allows episodic and semantic traces to be affected differently by events during the experiment. In the current version of eSAM, we have assumed that the episodic trace strengths are affected by experimental events and that the semantic trace strengths are not. This assumption seemed the most conservative and parsimonious to us. We do not rule out the possibility that semantic associations might be affected by events during the experiment, but any such mechanism would seem to add undue complexity to the model at this point.

Finally, the current version of eSAM assumes that semantic associations play a role only in retrieval. The model could have been structured to permit such associations to affect episodic processing during study. However, we took a more conservative and parsimonious approach, avoiding the addition of another mechanism absent a demonstrated need for it. Such an additional mechanism does not appear necessary to simulate these data. Nevertheless, we do not rule out the addition of such a mechanism as and when it proves necessary to model other behavioral data.

Measurement of Semantic Association Strength

Although LSA and WAS are both consistent with the SAM framework, we found that WAS provided a better fit to the category clustering data (see Simulation 2). This difference in fit seems attributable to the fact that, relative to LSA, WAS has both stronger associative coupling within natural categories and weaker associative coupling across different categories. In addition, whereas LSA strengths are based on co-occurrences in written text, WAS strengths are based on actual subjects' association protocols, thus more naturally assessing the tendency of one word to bring another to mind, in accord with the associative mechanism in eSAM.

The difference in fit using different semantic association metrics also suggests that some portion of the deviations from behavioral data may be attributable to measurement issues, rather than to features of the model itself. It seems likely that an even better fit of semantic data could be obtained using a semantic matrix consisting only of direct pairwise associations, as measured by word association norms, particularly to the extent that such norms reflect the actual asymmetric nature of word associations, in contrast to the symmetry imposed by both LSA and WAS. However, as of yet, direct associations have not been measured for a sufficiently large number of words as to permit exclusive use of such associations in a model such as eSAM, which requires a pairwise association strength value for each pair of words in a large lexicon.

Contextual Drift

As was noted previously, eSAM uses a simplified contextual drift mechanism that does not represent context as a vector of features and that does not change context after each time increment or item presentation (as in Howard & Kahana, 2002; Mensink & Raaijmakers, 1988; Murdock, 1997) but, rather, assumes that context changes after each list. However, our simplified implementation of context will likely require modification to permit eSAM to simulate other data. For one thing, the current implementation does not allow eSAM to retrieve items on a particular list from among several lists (e.g., Shiffrin, 1970), because there is no specific context marker for a particular list. In addition, it may prove important to distinguish between contexts associated with particular items and those associated with particular lists, as in the large literature on source memory judgments (see K. Mitchell & Johnson, 2000, for a review). To simulate these sorts of effects, eSAM could be modified to include a richer representation of context and, specifically, a mechanism for using items as retrieval cues for context, in addition to the reverse.

Scale Invariance of Recency and Temporal Contiguity

One shortcoming of dual-store models is their reliance on STM rehearsal to explain the recency effect and the lag-recency effect. As was noted previously, STM rehearsal can explain the recency effect in immediate free recall; it can also explain the elimination of the recency effect in delayed free recall by assuming that STM is emptied of studied items when a demanding distractor task is interposed between study and test (Postman & Phillips, 1965). However, the recency effect also occurs when a demanding distractor task is interposed between item presentations during study, as well as between study and test—a condition referred to as *continuous distractor* free recall (Bjork & Whitten, 1974). In this continuous distractor condition, overall recall is diminished, but the relative advantage accorded to end-of-list items is essentially the same as in immediate free recall. This *long-term recency effect* has been replicated many times, and it holds over a wide range of interpresentation intervals (Glenberg, Bradley, Kraus, & Renzaglia, 1983; Glenberg et al., 1980; Howard & Kahana, 1999). Dual-store models, such as SAM and eSAM, have difficulty explaining the long-term recency effect and its invariance across time scales, because the distractor task that displaces the contents of STM in delayed free recall should also displace the contents of STM in continuous distractor free recall.

Dual-store models also have difficulty explaining the finding that lag-recency is not disrupted in continuous distractor free recall (Howard & Kahana, 1999). As with the recency effect, STM rehearsal can explain the occurrence of the lag-recency effect in immediate free recall and in delayed free recall: Temporally contiguous items—those presented in nearby serial positions—would spend more time being jointly rehearsed together in STM than would items from more remote serial positions. Temporally contiguous items would, therefore, be more strongly associated episodically and more likely to be recalled in nearby output positions, due to cuing with interitem episodic associations. However, STM rehearsal cannot account for the occurrence of the lag-recency effect in continuous distractor free recall, because an interitem distractor should disrupt the formation of interitem episodic associations in STM. The scale invariance of the lag-recency effect therefore requires an alternative explanation.

One way to account for the long-term associative effects observed by Howard and Kahana (1999) involves a contextual drift mechanism that is somewhat different from that posited in eSAM, in Mensink and Raaijmakers (1988), and in Murdock (1997). Howard and Kahana (1999) proposed that recall of an item results in a partial reinstatement of the context that obtained when that item was studied. This *retrieved context* serves as a retrieval cue for other items with a similar context at study, which are most likely to be items from nearby serial positions, thus yielding the lag-recency effect. Howard and Kahana (1999) also posited that retrieval transitions are driven by the relative similarity between the temporal contexts of different list items, resulting in the scale invariance of the recency and lag-recency effects. These notions led to the development of the temporal context model (TCM; Howard & Kahana, 2002), a mathematical model that incorporates mechanisms for contextual drift and the formation of item-to-context associations and is able to capture these long-term recency and lag-recency effects. However, because that model lacks much of the machinery in SAM and eSAM, it cannot account for the detailed aspects of free recall data that we fit in our simulations. Future work aimed at integrating the key ideas in eSAM and TCM may result in a model that can account for many of the known features of episodic free recall, including, in particular, the effects of prior semantic and episodic learning.

CONCLUDING COMMENT

We have demonstrated that the venerable SAM model can be implemented in a way that includes a lexicon extending beyond the words on a single list and representations of semantic associations that are used to guide retrieval from LTM. Implementing these features allowed eSAM to simulate several effects of semantic associations on free recall. At the same time, we demonstrated that including a contextual drift mechanism, as well as specialized mechanisms involving items repeated on multiple study lists, permitted eSAM to simulate several effects of episodic experience on free recall. These features thus allow eSAM to go beyond recall of a single list to model phenomena that reflect the effects of prior episodic experience and semantic knowledge on free recall.

REFERENCES

Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review*, **79**, 97-123.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. New York: Cambridge University Press.

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology*, **6**, 173-189.

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

Clark, S. E. (1995). The generation effect and the modeling of associations in memory. *Memory & Cognition*, **23**, 442-455.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, **58**, 17-22.

Estes, W. K. (1955a). Statistical theory of distributional phenomena in learning. *Psychological Review*, **62**, 369-377.

Estes, W. K. (1955b). Statistical theory of spontaneous recovery and regression. *Psychological Review*, **62**, 145-154.

Estes, Z. (2003). Domain differences in the structure of artifactual and natural categories. *Memory & Cognition*, **31**, 199-214.

Frender, R., & Doubilet, P. (1974). More on measures of category clustering in free recall—although probably not the last word. *Psychological Bulletin*, **81**, 64-66.

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

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, **91**, 1-67.

Glanzer, M. (1976). Intonation grouping and related words in free recall. *Journal of Verbal Learning & Verbal Behavior*, **15**, 85-92.

Glanzer, M., Koppenaal, L., & Nelson, R. (1972). Effects of relations between words on short-term storage and long-term storage. *Journal of Verbal Learning & Verbal Behavior*, **11**, 403-416.

Glenberg, A. M. (1977). Influences of retrieval processes on the spacing effect in free recall. *Journal of Experimental Psychology: Human Learning & Memory*, **3**, 282-294.

Glenberg, A. M., Bradley, M. M., Kraus, T. A., & Renzaglia, G. J. (1983). Studies of the long-term recency effect: Support for a contextually guided retrieval theory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **9**, 231-255.

Glenberg, A. M., Bradley, M. M., Stevenson, J. A., Kraus, T. A., Tkachuk, M. J., Gretz, A. L., et al. (1980). A two-process account of long-term serial position effects. *Journal of Experimental Psychology: Human Learning & Memory*, **6**, 355-369.

Gronlund, S. D., & Shiffrin, R. M. (1986). Retrieval strategies in recall of natural categories and categorized lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 550-561.

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

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

Jenkins, J. J., & Russell, W. (1952). Associative clustering during recall. *Journal of Abnormal & Social Psychology*, **47**, 818-821.

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

Kahana, M. J., & Caplan, J. B. (2002). Associative asymmetry in probed recall of serial lists. *Memory & Cognition*, **30**, 841-849.

Kahana, M. J., & Howard, M. W. (2005). Spacing and lag effects in free recall of pure lists. *Psychonomic Bulletin & Review*, **12**, 159-164.

Kahana, M. J., Howard, M. W., Zaromb, F., & Wingfield, A. (2002). Age dissociates recency and lag-recency effects in free recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 530-540.

Kahana, M. J., & Wingfield, A. (2000). A functional relation between learning and organization in free recall. *Psychonomic Bulletin & Review*, **7**, 516-521.

Klein, K. A., Addis, K. M., & Kahana, M. J. (2005). A comparative analysis of serial and free recall. *Memory & Cognition*, **33**, 833-839.

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

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to latent semantic analysis. *Discourse Processes*, **25**, 259-284.

Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning & Verbal Behavior*, **2**, 1-21.

Mensink, G.-J., & Raaijmakers, J. G. W. (1988). A model for interference and forgetting. *Psychological Review*, **95**, 434-455.

Mitchell, K., & Johnson, M. (2000). Source monitoring: Attributing memories to sources. In E. Tulving & F. I. M. Craik (Eds.), *Oxford handbook of human memory* (pp. 179-185). Oxford: Oxford University Press.

Mitchell, M. (1996). *An introduction to genetic algorithms*. Cambridge, MA: MIT Press.

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, **64**, 482-488.

Murdock, B. B. (1974). *Human memory: Theory and data*. Potomac, MD: Erlbaum.

Murdock, B. B. (1997). Context and mediators in a theory of distributed associative memory (TODAM2). *Psychological Review*, **104**, 839-862.

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, & Computers*, **36**, 402-407.

Nelson, D. L., McKinney, V. M., Gee, N. R., & Janczura, G. A. (1998). Interpreting the influence of implicitly activated memories on recall and recognition. *Psychological Review*, **105**, 299-324.

Peterson, L. R. (1966). Short-term verbal memory and learning. *Psychological Review*, **73**, 193-207.

Peterson, L. R., Wampler, R., Kirkpatrick, M., & Saltzman, D. (1963). Effect of spacing presentations on retention of a paired associate over short intervals. *Journal of Experimental Psychology*, **66**, 206-209.

Phillips, J. L., Shiffrin, R. J., & Atkinson, R. C. (1967). The effects of list length on short-term memory. *Journal of Verbal Learning & Verbal Behavior*, **6**, 303-311.

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

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

Postman, L., & Underwood, B. J. (1973). Critical issues in interference theory. *Memory & Cognition*, **1**, 19-40.

Raaijmakers, J. G. W. (1979). *Retrieval from long-term store: A general theory and mathematical models*. Unpublished doctoral dissertation, University of Nijmegen.

Raaijmakers, J. G. W. (2003). Spacing and repetition effects in human memory: Application of the SAM model. *Cognitive Science*, **27**, 431-452.

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, pp. 207-262). New York: Academic Press.

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

Raskin, E., & Cook, S. W. (1937). The strength and direction of associations formed in the learning of nonsense syllables. *Journal of Experimental Psychology*, **20**, 381-395.

Roediger, H. L., III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 803-814.

Roediger, H. L., III, Payne, D. G., Gillespie, G. L., & Lean, D. S. (1982). Hypermnesia as determined by level of recall. *Journal of Verbal Learning & Verbal Behavior*, **21**, 635-655.

Romney, A. K., Brewer, D. D., & Batchelder, W. H. (1993). Predicting clustering from semantic structure. *Psychological Science*, **4**, 28-34.

Shiffrin, R. M. (1970). Forgetting: Trace erosion or retrieval failure? *Science*, **168**, 1601-1603.

Shiffrin, R. M., & Raaijmakers, J. (1992). The SAM retrieval model: A retrospective and prospective. In A. F. Healy, S. M. Kosslyn, & R. M. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes* (Vol. 2, pp. 69-86). Potomac, MD: Erlbaum.

Shiffrin, R. M., Ratcliff, R., & Clark, S. E. (1990). List-strength effect: II. Theoretical mechanisms. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 179-195.

Steyvers, M., Shiffrin, R. M., & Nelson, D. L. (2005). Word association spaces for predicting semantic similarity effects in episodic memory. In A. F. Healy (Ed.), *Experimental cognitive psychology and its applications* (pp. 237-249). Washington, DC: American Psychological Association.

Tulving, E. (1966). Subjective organization and effects of repetition in multi-trial free-recall learning. *Journal of Verbal Learning & Verbal Behavior*, **5**, 193-197.

Tulving, E. (1968). Theoretical issues in free recall. In T. R. Dixon &

D. L. Horton (Eds.), *Verbal behavior and general behavior theory* (pp. 2-36). Englewood Cliffs, NJ: Prentice-Hall.

Ward, G., Woodward, G., Stevens, A., & Stinson, C. (2003). Using overt rehearsals to explain word frequency effects in free recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **29**, 186-210.

Wingfield, A., & Kahana, M. J. (2002). The dynamics of memory retrieval in older adults. *Canadian Journal of Experimental Psychology*, **56**, 187-199.

Wingfield, A., Lindfield, K. C., & Kahana, M. J. (1998). Adult age differences in the temporal characteristics of category free recall. *Psychology & Aging*, **13**, 256-266.

Zaromb, F., Howard, M. W., Dolan, E., Sirotin, Y. B., Wingfield, A., & Kahana, M. J. (2005). *Temporal associations and prior-list intrusions in free recall*. Manuscript submitted for publication.

(Manuscript received May 29, 2003;
revision accepted for publication April 5, 2005.)