Journal of Experimental Psychology:
Learning, Memory, and Cognition
2005, Vol. 31, No. 5, 933–953

0278-7393/05/$12.00 DOI: 10.1037/0278-7393.31.5.933

Theoretical Correlations and Measured Correlations: Relating Recognition and Recall in Four Distributed Memory Models

Michael J. Kahana
University of Pennsylvania

Daniel S. Rizzuto
California Institute of Technology

Abraham R. Schneider
New York University

This article addresses the relation between item recognition and associative (cued) recall. Going beyond measures of performance on each task, the analysis focuses on the degree to which the contingency between successful recognition and successful recall of a studied item reflects the commonality of memory processes underlying the recognition and recall tasks. Specifically, 4 classes of distributed memory models are assessed for their ability to account for the relatively invariant correlation ($\approx$ .5) between successive recognition and recall. Basic versions of each model either under- or overpredict the intertask correlation. Introducing variability in goodness-of-encoding and response criteria, as well as output encoding, enabled all 4 models to reproduce the moderate intertask correlation and the increase in correlation observed in 2 mixed-list experiments. This model-based analysis provides a general theoretical framework for interpreting contingencies between successive memory tests.

Keywords: recognition, recall, correlation, memory, modeling

This article addresses the relation between item recognition and associative, or cued, recall. The item-recognition task asks subjects to judge whether a target item was on a just-presented list; the cued-recall task asks subjects to generate the target paired with a given cue item. Recognition and recall are considered to be the two canonical episodic memory tasks, with both tasks requiring that subjects retain information about the item's presence within a temporally defined set (e.g., Tulving, 1983). Because a growing number of memory models now provide a unified theoretical framework for modeling both item recognition and associative recall, we asked whether these models can account for data on the contingency between successful recognition and successful recall of a given studied item. More generally, we examine how item-level correlations between successive memory tasks inform the key assumptions governing the representation, storage, and retrieval of information in those tasks.

Michael J. Kahana, Department of Psychology, University of Pennsylvania; Daniel S. Rizzuto, Biology Department, California Institute of Technology; Abraham R. Schneider, Center for Neural Science, New York University.

We acknowledge support from National Institutes of Health Grant MH55687 and National Science and Engineering Research Council of Canada Grant APA 146. Experiment 1 was carried out in Ben Murdock's lab at the University of Toronto in 1993. We thank Ben Murdock and Endel Tulving for stimulating our interest in this question and Bill Estes, Rich Shiffrin, Tom Wickens, Jim Neely, Bob Greene, Colin MacLeod, Dan Kimball, Marc Howard, and Jeremy Caplan for their insightful comments on a draft of this article.

Correspondence concerning this article should be addressed to Michael J. Kahana, Department of Psychology, University of Pennsylvania, Suite 303C, 3401 Walnut Street, Philadelphia, PA 19104. E-mail: kahana@sas.upenn.edu

Strength theory provided the earliest account of the relation between recognition and recall. According to this view, the study of the items on a list strengthens associations between each of the list items and some representation of the list itself. Outcomes of both recognition and recall tests depend on the strength of this item-to-list association. Consistent with the observation that recognition is usually easier than recall, this view held that the essential difference between these tasks is that recognition could be successfully performed with weaker associations than recall.

Despite its heuristic value, strength theory offered an overly simplistic view of recognition–recall differences. This theory was abandoned after researchers found that experimental variables had opposing effects on recognition and recall. For example, as compared with common words, rare words are easily recognized as having been presented in a recent list but are more difficult to recall (Gregg, 1976; Kinsbourne & George, 1974; MacLeod & Kampe, 1996). Other examples of dissociations include the effects of intentionality on memory encoding (Glenberg & Bradley, 1979; Schwartz & Humphreys, 1974), the effects of context change on memory retrieval, and the effects of damage to the medial-temporal lobe (Hirst, 1986; Vargha-Khadem et al., 1997), at least when interitem similarity is not high (Holdstock et al., 2002).

Generate–recognize theory provided an alternative conception of the differences between recognition and recall (e.g., Bahrick, 1970). According to this view, recall involves two stages: subjects first generate possible responses, and then apply a recognition test to decide whether any of the generated responses were on the list. The recognition task differs from recall in that the generate stage is absent. A strong version of this model predicts that recallable items will always be recognized. Contrary to this prediction, Tulving and colleagues (e.g., Tulving, 1968; Tulving & Thompson, 1973) found that unrecognized items may often be subsequently recalled when prompted by an appropriate retrieval cue.

Contingency Analysis of the Recognition–Recall Relation

In studying the relation between recognition and recall, Tulving and Thompson (1973) adopted the successive testing technique, a technique that had been used previously to examine one-trial associative learning and to assess the role of associative unlearning (see Kahana, 2000, for a review). In Tulving and Thompson's procedure, subjects studied a list of A–B word pairs and were then tested successively: first by item recognition and then by cued recall. In the item-recognition test, subjects saw B items from each of the studied pairs intermixed with nonlist items. Subjects responded "yes" to items if they remembered seeing them in the study list. In the cued-recall test, subjects attempted to recall the B items when given the A items as cues. In this manner, memory for each of the B items was tested twice—first by recognition and then later by recall. Tulving and colleagues observed that some items that subjects failed to recognize as having been presented in the study list were nonetheless correctly recalled on the cued-recall test—a finding they referred to as the "recognition failure of recallable words."

Tulving and Wiseman (1975) measured the statistical contingency (or association) between item recognition and cued recall at the level of individual items. Their analysis revealed a simple relation between the conditional probability of recognition given recall, $P(R|C)$, and the probability of recognition itself, $P(R)$. This relation, known as the Tulving–Wiseman function, describes a moderate degree of dependency between item recognition and cued recall. Expressed using Yule's Q—a measure of association for 2×2 contingency tables—one finds values ranging from .45 to .65 across a wide range of experimental conditions (Kahana, 2000; Nilsson & Gardiner, 1991).[1] Although higher recognition–recall dependencies have been observed as a result of shallow encoding or semantic redundancy of study pairs (Nilsson & Gardiner, 1993; Nilsson, Law, & Tulving, 1988), the basic finding of moderate dependence is robust.

The moderate degree of dependence between recognition and recall contrasts with other task comparisons, particularly those involving comparisons of implicit and explicit memory tasks. Independence ($Q \sim 0$) is typically observed when subjects perform successive recognition and fragment completion or successive fragment completion tasks given with implicit instructions on one test and explicit instructions on the other (Hayman & Tulving, 1989a, 1989b; Tulving, Schacter, & Stark, 1982). These findings of independence have led to very strong claims concerning the existence of separate memory systems supporting implicit and explicit memory (Tulving & Schacter, 1991).

Intertask correlations can also be very high, as in cases in which the same information is probed on two successive tests. For example, Kahana (2002) had subjects study a list of 12 A–B pairs. All the studied pairs were then tested twice, once in each of two test phases (designated as Test 1 and Test 2). On Test 1, half of the studied pairs were cued in the forward order, and the other half were cued in the backward order. On Test 2 (which was given after a brief delay), half of the pairs were tested in the same order as in Test 1 and half were tested in the reverse order. The correlation between successful recall of pairs tested in an identical manner on the two retrieval occasions was .99. More surprisingly, the correlation was .97 for pairs tested in the forward order on one of the tests and the backward order on the other (e.g., A–? on Test 1 and ?–B on Test 2). Kahana interpreted the similarity of these correlations as evidence for the associative symmetry hypothesis (Asch & Ebenholtz, 1962), which sees A–B associations as a holistic conjunction of both A and B items rather than independent forward and backward links (e.g., Wolford, 1971).

The preceding examples illustrate that the moderate correlation between successive recognition and recall must say something about the information processing underlying these two memory tasks. Indeed, the moderate correlation between item recognition and cued recall, coupled with numerous experimental dissociations in which manipulated variables selectively influence either recognition or recall, led theorists to advocate for a distinction between item-specific and relational (or associative) information (Humphreys, 1978; Murdock, 1974)—a distinction that became formalized in a host of computational memory models (Gillund & Shiffrin, 1984; Hintzman, 1988; Humphreys, Pike, Bain, & Tehan, 1989; Mensink & Raaijmakers, 1988; Metcalfe, 1985; Metcalfe, 1985; Murdock, 1982, 1997; Norman & O'Reilly, 2003). Although these models have been used extensively to account for recognition and recall data individually, the basic assumptions about the dependence or independence of the information supporting these tasks have not been carefully evaluated. However, before one can evaluate these assumptions, one must consider the complexities involved in interpreting correlations between successive tests.

Our overarching goal is to provide a rigorous foundation for the interpretation of intertask correlations and to use this approach to help us understand how correlations between memory tasks can help constrain models of those tasks. Our focus on the recognition–recall correlation is fitting because recognition and recall are the two most widely studied memory tasks, and because major memory models have been developed to provide a common theoretical framework for their analysis. Furthermore, theories of recognition and recall have been specified in sufficient detail that we can derive explicit predictions about their correlation. Finally, there are extensive data demonstrating the moderate correlation between successive recognition and recall tasks, with these correlations being consistent across studies that vary widely in their methodology and in the overall level of recognition and recall performance.

This article is organized as follows. First, we present a statistical analysis of factors that may influence the correlation between any two successive memory tests. Next, we introduce four distributed memory models (DMMs) that have been used to model item recognition and associative (cued) recall. These models are of particular interest because they make representational assumptions that allow one to derive, analytically, the recognition–recall correlation. These derivations are then presented for simplified versions of each of the four models. Using Monte Carlo methods, we are able to simulate extended versions of each model and show that variability at encoding and variability at retrieval play crucial roles in determining intertask correlations. The latter part of the article reports two experiments that test and confirm these predictions,

[1] Yule's Q can take on values ranging from -1.0 (perfect negative correlation) to $+1.0$ (perfect positive correlation).

thus demonstrating the important role played by variability in modeling the correlation between successive memory tasks.

Factors Affecting the Correlation Between Successive Tests

Because of problems noted in the graphical analysis of Tulving and Wiseman (Hintzman, 1992), we approach the general problem of intertask correlations by directly examining the 2 × 2 contingency table. This can be done by fitting models to either (a) the probabilities of successful outcomes on Tests 1 and 2 as well as the correlation (Q) between these outcomes or (b) the probability of success on Test 1, the probability of success on Test 2 conditional on Test 1 success, and the probability of success on Test 2 conditional on Test 1 failure. We use the former approach; the latter approach has been used effectively by Humphreys and Bowyer (1980) and Batchelder and Riefer (1995).

Simpson's paradox poses a challenge for the interpretation of contingency tables. This paradox refers to the fact that collapsing data across subjects or items can give rise to relations that were not present in the precollapsed data (e.g., Hintzman, 1981; Hintzman & Hartry, 1990). Consider for example what would happen if some items attract a good deal of attention and are thus very well encoded during the study phase, whereas other items attract little attention and are thus very poorly encoded during the study phase. Collapsing across these two classes of items could produce a high correlation between recognition and recall even if the correlation within each class of items was rather low. Although it could be argued that such effects render the analysis of intertask correlations meaningless, we show how these complications can be taken into account in a model-based analysis of the correlations between successive tasks.

Figure 1 shows a causal diagram depicting the effects of different sources of variability on the observed correlation between item

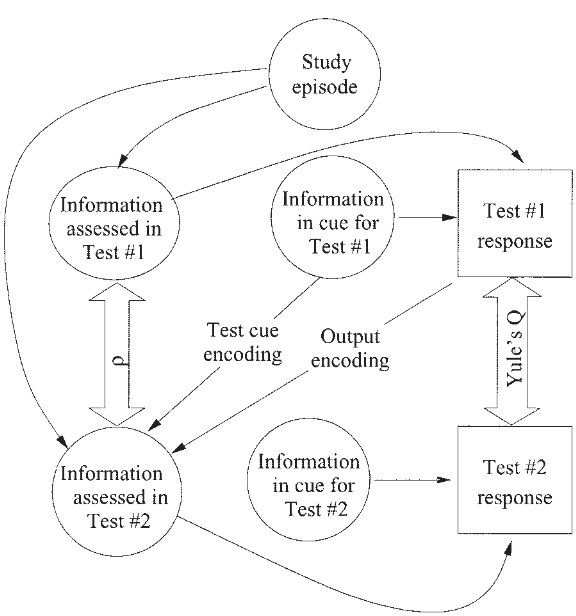

Figure 1. Causal model illustrating the factors that can influence the observed correlation between item recognition and cued recall.

recognition and cued recall. Consider the study episode: For a given subject, each pair will be encoded in an unpredictable manner. Fluctuations in attention, subject-specific coding of word pairs, and differences in mnemonic strategies produce variability that can be unique to each studied pair. This variability in *goodness of encoding,* which is difficult to estimate or control (cf. Hintzman & Hartry, 1990), can induce a positive correlation between successive memory tasks (Flexser & Tulving, 1978; Hintzman, 1987; Kahana, 2000). Consider what would happen if subjects, faced with a very long list, chose to attend to the first few word pairs and rehearse those pairs while ignoring the remaining list items. Because the first few pairs will be remembered very well and the remaining (ignored) pairs will be neither recognized nor recalled, we will find that $Q \approx 1$, even if the inherent correlation between information driving recognition and recall is very small.

Variability in goodness of encoding is not the only factor that can influence the correlation between successive tests. Because retrieval is cue dependent, encoding of the cue item is also crucial for memory performance. Consider the case of identical successive memory tests (e.g., recall of B given A on two successive tests). Suppose, further, that there is no learning in Test 1 and no forgetting in the interval between Tests 1 and 2. One might expect that items correctly remembered in Test 1 will also be correctly remembered on Test 2; conversely, items that are not remembered on Test 1 will not be remembered on Test 2. This assumes that the cue item (A_i) is encoded identically on Test 1 and Test 2, and that the criterion for responding is identical at both tests. However, if the retrieval cue is encoded more effectively on Test 1 than on Test 2, or if the criterion for recall is lowered, items that are correctly remembered on Test 1 may not be remembered on Test 2, thereby reducing the intertask correlation. If two variables are correlated, but independent sources of variability affect each one, their correlation will be attenuated.

A final complication comes from the influence of Test 1 on Test 2. On each test trial, subjects store information about the test cue and the information it retrieves (see Figure 1). We refer to this storage as *output encoding.* Humphreys and Bowyer (1980) examined whether output encoding during item recognition facilitated subsequent cued recall. Subjects studied pairs of words and then performed successive recognition and recall tasks. Some items tested during recall were not presented during the prior recognition test. Humphreys and Bowyer found higher cued-recall performance for items that were previously tested as compared with those that were not. This suggests that output encoding during item recognition facilitates subsequent cued recall.

Humphreys and Bowyer (1980) hypothesized that output encoding may alter the observed correlation between item recognition and subsequent cued recall. If subjects store more relevant information for recognized than for nonrecognized items, then recall should be higher for recognized items, thus increasing the correlation between successive tests. If output encoding is identical for recognized and nonrecognized items, one would not expect to find such an effect.

Humphreys and Bowyer (1980) presented evidence that although recognition facilitates later recall even when an item is not recognized (Begg, 1979; Donnelly, 1988), the boost to recall is greater for recognized than for nonrecognized items. This enhanced output encoding for recognized items leads to an increase in the correlation between recognition and subsequent recall. Some experiments (e.g., Wiseman & Tulving, 1976) fail to show significant output encoding but nonetheless yield moderate dependency between item recognition and cued recall. This suggests that output encoding only partially contributes to the observed correlation between recognition and subsequent recall.

In summary, variability in storage should increase the correlation between tasks so long as the tasks interrogate the same episode in memory. Variability in the retrieval process will lower the correlation. Output encoding can increase the correlation between tasks under specific assumptions about the nature of the information being stored during Test 1. This analysis holds for any pair of successive memory tasks, but we focus our attention on successive item recognition and cued recall.

Model-based analyses that incorporate these factors can aid our understanding of the recognition–recall relation. We apply distributed memory models to data on recognition, recall, and their contingency relations. For simplified versions of these models, one can derive analytic expressions for the base level of correlation between recognition and recall. More complete versions of these models will include other factors that can modulate these base correlations, as shown in Figure 1. Monte Carlo techniques can then be used to study the behavior of these more realistic, but less analytically tractable, models.

Rather than focusing on a single model, we examine a portfolio of four models that differ along two dimensions: (a) their associative mechanism and (b) the comparison process they use in item recognition. The next section describes the basic machinery underlying these models.

DMM

DMMs assume that the stream of incoming experience is parsed into meaningful units. Each unit is then represented by a set of abstract feature values; mathematically, this set describes a vector in a high-dimensional feature space. These distributed representations are then stored in a single memory system (a composite representation containing all of the stored items and associations). There are many ways to store these representations. The three basic types of storage used by the DMMs considered in this article are *autoassociation, heteroassociation,* and *direct storage.*

An autoassociative mechanism binds features in such a way that a part can be used to retrieve the whole (*redintegration*). A heteroassociative mechanism binds features in such a way that one activated pattern can be used to retrieve another. Two mathematical operations have been proposed to form autoassociative and heteroassociative memory representations. In one approach (Anderson, Silverstein, Ritz, & Jones, 1977; Pike, 1984) the association is formed by taking the outer product of two N-dimensional vectors. The result of this operation is an $N \times N$ matrix (see Jordan, 1986, for details). In a second approach (Murdock, 1979), the association is formed by taking the vector convolution of two N-dimensional vectors. The result of this operation is a $2N - 1$ dimensional vector.[2] In both cases, if the vectors being associated are identical, the operation is autoassociation. If the vectors being associated are different, the operation is heteroassociation. Once these associative representations are formed, they can be added to a single memory structure. The same structure can represent many different associations.[3]

A different mechanism for storing information in memory, termed *direct storage,* adds vectors representing units of information directly to the memory structure—without forming any kind of associations. Anderson (1973) demonstrated that a model based on direct storage could account for a wide range of data on single-item recognition. This type of storage, however, would not allow for associative retrieval or pattern completion (Weber & Murdock, 1989).

Murdock (1982) demonstrated how a single model combining both direct storage and heteroassociative mechanisms could account for the data on item recognition and cued recall. Metcalfe (1985) modeled item recognition and cued recall by combining autoassociative and heteroassociative storage within a single model. Murdock and Metcalfe's models both use convolution as the heteroassociative storage mechanism. Humphreys, Bain, and Pike (1989) proposed a matrix model for both recognition and recall. This model used direct storage for encoding items, and used matrix multiplication as the heteroassociative mechanism.

Modeling Recognition and Recall

Our four DMMs represent a factorial combination of two kinds of associative mechanisms (convolution and matrix multiplication) and two approaches to modeling the recognition process (autoassociation and direct storage). Because all four models use similar recall mechanisms, we focus first on the two different approaches used to model the recognition process.

Local- Versus Global-Match Recognition Models

For models using direct storage (summation), a simple comparison or match of the probe item with the contents of memory provides information about whether that item was present in the list. This comparison is termed a *global match* because the value or strength of the match reflects the contributions of all the list items. Global-match models assume that the recognition decision is based on the summed similarity of the probe item with all of the traces stored in memory (see Clark & Gronlund, 1996, for a review).

Autoassociative models (e.g., Metcalfe, 1985, 1991; Norman & O'Reilly, 2003) approach the recognition problem in a different way. During study, each list item is autoassociated and then these autoassociations are stored in memory. Recognition is a two-stage process. First, the probe item is used to retrieve the associated information in memory. If the probe item was one of the items on the list, the retrieved information will include the probe item. In the second stage, the retrieved information is matched against the probe item itself. If this local match exceeds a fixed criterion, the model returns a positive response. We refer to this two-stage model as a *local-match model* because the probe item is compared

[2] The convolution of two vectors, **f** and **g** is defined by the equation $(f * g)_m = \sum_i f_i g_{m-i}$, where m is the index to the elements in the convolution vector and i indexes the elements in the item vectors **f** and **g**. The asterisk (∗) denotes the convolution operator.

[3] Biologically inspired nonlinear networks can also perform autoassociation and heteroassociation (e.g., Chappell & Humphreys, 1994; Lisman, Jensen, & Kahana, 2001). We analyze linear DMMs because analytic expressions for their behavior can be easily obtained (Weber, 1988).

Table 1
Factorial Analysis of the Storage Processes in Four Classes of Distributed Memory Models

Recognition process	Matrix product	Vector convolution
Local match	$W_k = W_{k-1} + (\mathbf{f}_k + \mathbf{g}_k)(\mathbf{f}_k + \mathbf{g}_k)'$ $= W_{k-1} + \mathbf{f}_k\mathbf{g}_k' + \mathbf{g}_k\mathbf{f}_k' + \mathbf{f}_k\mathbf{f}_k' + \mathbf{g}_k\mathbf{g}_k'$	$\mathbf{m}_k = \mathbf{m}_{k-1} + (\mathbf{f}_k + \mathbf{g}_k) * (\mathbf{f}_k + \mathbf{g}_k)$ $= \mathbf{m}_{k-1} + 2\mathbf{f}_k * \mathbf{g}_k + \mathbf{f}_k * \mathbf{f}_k + \mathbf{g}_k * \mathbf{g}_k$
Global match	$W_k = W_{k-1} + \mathbf{f}_k\mathbf{g}_k' + \mathbf{g}_k\mathbf{f}_k' + \mathbf{f}_k\mathbf{r}' + \mathbf{r}\mathbf{g}_k'$	$\mathbf{m}_k = \mathbf{m}_{k-1} + \mathbf{f}_k + \mathbf{g}_k + \mathbf{f}_k * \mathbf{g}_k$

with the retrieved information as opposed to the entire contents of memory.

Recent work has highlighted the limitations of simple global- and local-match models of recognition, advocating instead for a dual-process approach in which local- and global-match processes operate in tandem (Norman & O'Reilly, 2003). While recognizing that such a hybrid approach is likely to provide a more successful account of recognition memory phenomena, our goal of contrasting distinct classes of models is served by setting up a strong contrast. We have therefore chosen to analyze the process-pure global- and local-match models separately. A hybrid model is likely to generate predictions that fall between those of the process-pure variants.

Modeling Cued Recall

In cued recall, subjects study a list of word pairs denoted F_1–G_1, F_2–G_2, . . . , F_L–G_L, where L represents the number of pairs in the list. At test, the experimenter cues with each F item for recall of the corresponding G item. In the models, cued recall depends on the storage of the heteroassociation of the vectors representing F and G. Retrieval involves cuing the memory with the probe item and applying an associative retrieval operation to the memory system. In the matrix models, multiplying the probe with the memory matrix yields the retrieved vector. In the convolution models, correlating the probe item vector with the memory vector yields the retrieved vector.[4] The probability of retrieving a given target is assumed to be proportional to the match of the retrieved vector with the target vectors.[5] In all of the models presented here, the information used for both recognition and recall is assumed to reside in a common episodic memory store.

Mathematical Characterization of the Models

Table 1 gives the storage equations for the four DMMs considered here. In each equation, bold lowercase characters represent vectors and capital letters represent matrices (**m** denotes the memory vector in the convolution–correlation models, W denotes the weight matrix in the matrix models), $\mathbf{f}_k$ and $\mathbf{g}_k$ represent the studied items, the subscript k indexes the current pair being stored, a prime is used to denote the transpose of a vector (e.g., $\mathbf{f}'$), and **r** represents a fixed vector of unit length.

All four models will recognize an item, **f,** if the strength of the information signaling the item's presence in memory, R, exceeds a decision criterion. In the global-match convolution–correlation model, $R = \mathbf{f} \cdot \mathbf{m}$ (for a more sophisticated treatment of the decision process, see Hockley & Murdock, 1987). In the local-match convolution–correlation model, $R = (\mathbf{f}\#\mathbf{m}) \cdot \mathbf{f}$. In the global-match matrix model (e.g., Humphreys, Pike, et al., 1989, without context), $R = \mathbf{fr}' \cdot W$. In the local-match matrix model, $R = W\mathbf{f} \cdot \mathbf{f}$ (cf. Rizzuto & Kahana, 2001). In all four models, recall performance is proportional to the match between the retrieved and desired information, denoted C. In the convolution–correlation memory models, $C = (\mathbf{f}\#\mathbf{m}) \cdot \mathbf{g}$, where **g** is the desired target item. In the matrix memory models, $C = W\mathbf{f} \cdot \mathbf{g}$. See Appendix A for a more detailed description of the four models.

Mathematical Analysis of the Recognition–Recall Correlation

Our goal in this section is to use models of recognition and recall to derive predictions for the correlation between these tasks at the level of the individual items.[6] With variables that represent the quality of information driving item recognition, R, and cued recall, C, the theoretical correlation between R and C is given by

$$\rho_{RC} = \frac{\mathrm{cov}(R, C)}{\sqrt{\mathrm{var}(R)\mathrm{var}(C)}}.$$

Table 2 gives the variance and covariance expressions used to calculate the theoretical correlation between recognition and recall in each of the four models (detailed model descriptions and derivations of these expressions can be found in Appendix A). In deriving these analytic expressions, we used base versions of each model; versions that do not take account of variability in goodness of encoding, variability in the retrieval process, output encoding, or interitem similarity. In the next section, we examine these complicating factors using Monte Carlo simulation methods, but first let us consider the implications of these basic implementations of each model.

In all four models, both the number of features, N, and the list length, L, contribute to the variance of R and C. This is because adding correlated features to the memory matrix or vector will contribute to the variance of the matching values. Because N must be relatively large to support recognition and recall performance, the higher order terms do not contribute substantially to the cor-

[4] Correlation is an approximate inverse of convolution. **f** correlated with **g** is defined by the equation $(f\#g)_m = \sum_i f_i g_{m+i}$, where the pound sign (#) denotes the correlation operator.

[5] The mapping between retrieved and target vector, which is sometimes called *deblurring,* can also be achieved using a more neurally plausible dynamical rule (Anderson et al., 1977; Farrell & Lewandowsky, 2002).

[6] Although it would be better to derive Q directly, this would entail working with integrals of the normal distribution function, and thus would not allow for closed form expressions for the correlation. As long as sphericity is approximately satisfied by the contingency data, the product–moment correlation, ρ, should be nearly equal to Q.

Table 2
Variance and Covariance of Matching Strengths Driving Item Recognition and Recall Performance

Model	var(R)	var(C)	cov(R, C)
Global-match convolution	$(2.75L + 2)N^{-1} + N^{-2} + 0.25LN^{-3}$	$(2.17L + 6.67)N^{-1} + 4N^{-2} + (0.83L + 0.33)N^{-3}$	$3.75N^{-1} + N^{-2} + 0.25N$
Local-match convolution	$\frac{1}{6}[(40L + 232)N^{-1} + (27L + 276)N^{-2} + (20L + 80)N^{-3} + (27L + 12)N^{-4}]$	$\frac{1}{3}[(10L + 38)N^{-1} + (3L - 18)N^{-2} + (5L - 73)N^{-3} + 3LN^{-4}]$	$\frac{1}{3}[50N^{-1} + (15L + 90)N^{-2} + (12L - 35)N^{-3} + (3L - 3)N^{-4}]$
Global-match matrix	$(L + 2)N + (L + 2) + (2L + 5)N^{-1} + (2L + 6)N^{-2}$	$(2L + 6)N^{-1} + (2L + 18)N^{-2} + (2L + 6)N^{-3}$	$2 + 4N^{-1} + 4N^{-2}$
Local-match matrix	$12N^{-1} + (24L + 50)N^{-2} + (8L^2 + 40 + 54)N^{-3}$	$8N^{-1} + (12L + 22)N^{-2} + (4L^2 + 12L + 22)N^{-3}$	$8N^{-1} + (8L + 30)N^{-2} + (12L + 30)N^{-3}$

relation. For the most part, the ratio of the lowest order term in the covariance to the geometric average of the equivalent order terms in the variances will drive the recognition–recall correlation.

Figure 2 plots the correlation between R and C as a function of L for each of the models. In both the global- and local-match convolution models, the dominant terms in the variances and covariance are of the order N^{-1}. However, because L contributes to the variance but not to the covariance, these models predict that the correlation will go to zero as L grows large. This is shown in the upper left and upper right panels of Figure 2. In the global-match matrix model (lower left panel of Figure 2), the dominant term in the covariance is 2, and the dominant term in the product of the variances is $2L^2$. Here again, the correlation approaches zero as L grows large.

In the local-match matrix model, the dominant terms in the variances and covariances are all of the order N^{-1}. However, unlike the other models, this term does not depend on L in either the covariance or the variance expressions. For this reason, this model exhibits a high correlation even for large values of L (lower right panel of Figure 2). For small N, the higher order terms in the variance contribute to the correlation, and because these terms depend on L, the correlation does decrease. For high N, however, the local-match matrix model predicts very high correlations for all values of L.

For typical experimental parameters, only the local-match matrix model produces a substantial positive correlation between recognition and recall. In the other three models, the positive covariance between recognition and recall is a negligible fraction of the variance. Consequently, these three models predict a near-zero correlation between item recognition and cued recall for the kinds of experiments that typically yield moderate correlations ($Q \approx .5$).[7]

Simulating Encoding Variability and Output Encoding

In deriving the correlation between R and C, we made a number of simplifying assumptions. First, we assumed that subjects store a constant amount of information for each studied pair. Second, we assumed that both the recognition and recall cues were coded perfectly. Third, we assumed that both Test 1 and Test 2 assess the information stored during the initial study phase. As discussed previously, Test 2 assesses information stored during study as well as any additional information encoded during Test 1 (see Figure 1). In this section, we use Monte Carlo simulations to evaluate generalized versions of each of the four DMMs, considering in particular the effects of encoding variability and output encoding on the recognition–recall correlation.

Simulation 1: Encoding Variability

Here we consider the effect of variability in goodness of encoding on the recognition–recall correlation. Probabilistic encoding of the constituent features of an item can be used to vary the quality of encoding of items and associations (e.g., Murdock, 1997; Murdock & Lamon, 1988; Shiffrin & Steyvers, 1997). Following this approach, a single presentation of a stimulus results in a sample of its features being encoded. Each feature is encoded with probability p and not encoded (set to zero) with probability $1 - p$. This formalizes the notion of goodness of encoding: Better encoding translates into higher p values, worse encoding translates into lower p values.

Hockley and Cristi (1996) have shown that stressing item encoding does not facilitate memory for associations; however, stressing associative encoding boosts both memory for item and associative information. Because cued recall stresses associative encoding, we limit our analysis to variability in the goodness of encoding for the studied pair as a whole. Pairs that are encoded well (i.e., p is large) will have a higher probability of recognition and recall than pairs that are encoded poorly (i.e., p is small), and variability in p will increase the correlation between successive tasks. If the variance in p is large enough, the correlation between recognition and recall should approach unity, even if item and associative information are independent.

Method. We simulated the successive study and test of 32 pairs of N-dimensional item vectors, denoted $\mathbf{f}_i$, $\mathbf{g}_i$, whose elements are independent and identically distributed random variables, $f_i(k) \sim \mathcal{N}(0, \sqrt{1/N})$. All of the study items served as competitors in the recall phase. During study, items were encoded probabilistically, with the value of p drawn separately for

[7] Metcalfe (1991) suggested that local- but not global-match convolution models can account for the moderate recognition–recall dependency. When the two models simulated the study of three word pairs, Metcalfe found that the local-match version produced a moderate correlation between item recognition and cued recall, whereas the global-match version predicted independence. As shown in Table 2 of the present article, the behavior of these model depends on list length and vector dimensionality. When simulating lists of just three pairs, and using vectors of low dimensionality, as in Metcalfe, one observes a somewhat higher correlation for the local-match model. Both models, however, predict near independence when list length is increased to values used in experimental studies.

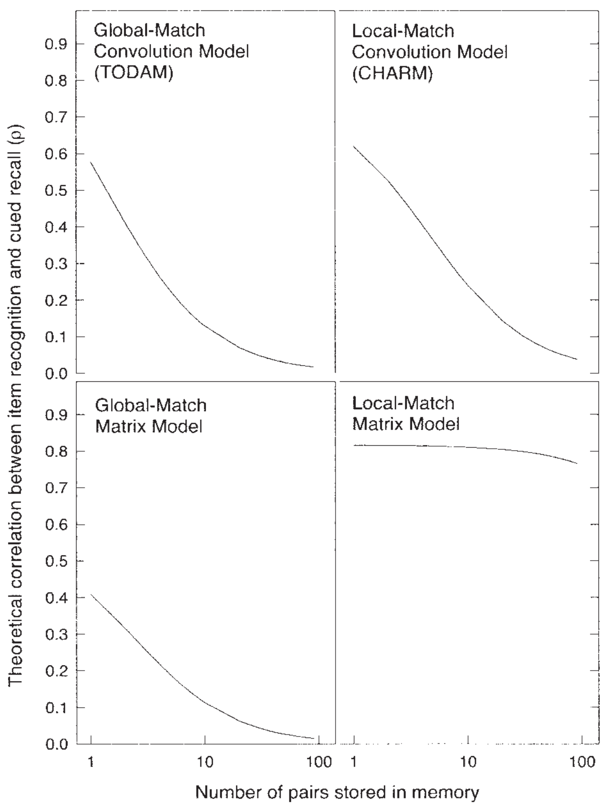

Figure 2. Theoretical correlation between recognition and recall for four distributed memory models: local- and global-match models that use either convolution or matrix operations to associate item vectors. In each of the models, the derived correlation depends on the number of pairs stored in memory. The dimensionality of the item vectors (N) was set to 1,000. TODAM = theory of distributed associative memory; CHARM = composite holographic associative recall model.

each pair of items, $p \sim \mathcal{N}(\mu_p, \sigma_p;$ the same value of p was used for the two members of a given pair).

To examine the effect of variability in p on Q, we varied σ_p between 0 and .25. In both recognition and recall, the test probes were encoded probabilistically (with values of .8, .9, and 1.0), but were not stored in memory. Other parameters were μ_p, N, the resemblance criterion for recognition, and the resemblance criterion for recall. These values were adjusted to achieve a recall probability of approximately .35 and a hit rate of approximately .75 for each of the four models. We examine the effect of output encoding in the next simulation.

Results. As shown in Figure 3, Q rises as σ_p increases in each of the four models and for all levels of the encoding of the test probes. Variability in p raises the correlation between tasks because both tasks assess memory for the same study episode. As indicated in the analytic solution for the correlation (Figure 2), without variability in p, only the local-match matrix model predicts a high correlation between recognition and recall.

Simulation 2: Output Encoding

Here we consider the effect of output encoding during recognition on subsequent recall. During a recognition test, information evoked by a given test probe will depend on whether the probe is recognized as an item experienced on the list. Subsequent cued recall will benefit if the information stored during recognition is correlated with the information tested in recall. Models that assume independence of item recognition and cued recall should not predict significant facilitation of recall. For models that assume some correlation between recognition and recall, facilitation of recall should occur only if the storage of information used in recall is greater for items recognized as members of the study list. In this case, one would also predict an increase in the correlation between successive tests.

We simulated the effect of output encoding on the recognition–recall relation in each of our four models. Local- and global-match models fundamentally differ in their implementation of the recognition process. The local-match models first use the recognition probe to retrieve the contents of memory. They judge an item as "old" if the retrieved information matches the probe. For local-match models, recognizing an item as old should lead to the encoding of the retrieved item and associative information. Global-match models base recognition judgments solely on the match of the probe item with the contents of memory. For these models, judging an item as old should just lead to the encoding of the probe

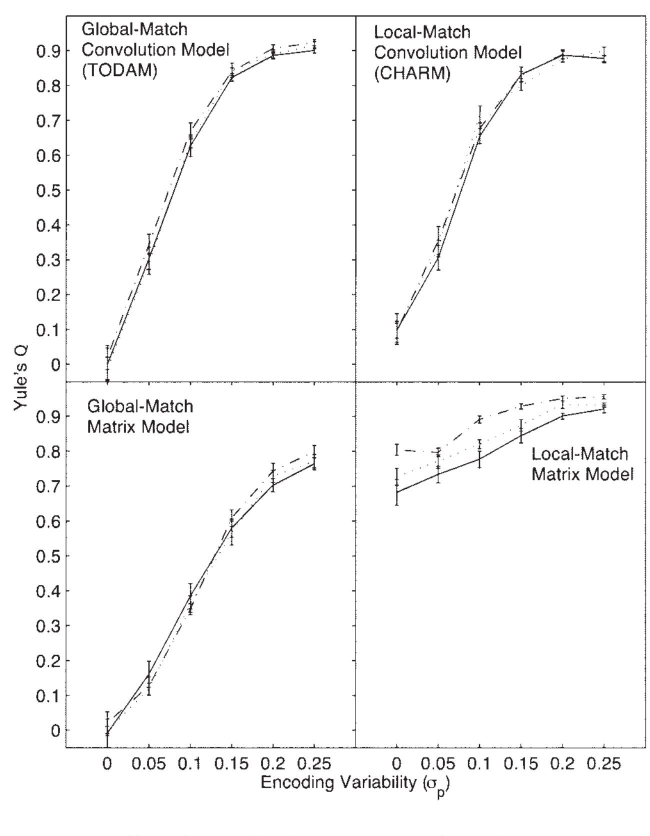

Figure 3. Effect of variability in the goodness of encoding on the dependency between item recognition and cued recall. Each panel shows results for a different distributed memory model. Q is plotted as a function of the standard deviation of the probabilistic encoding parameter, p, for the A–B pairs ($0 < p < 1$). The solid line represents a probabilistic encoding level of .8; dotted line, .9; and dot-dashed line, 1.0. Error bars represent standard error. See text for details. TODAM = theory of distributed associative memory; CHARM = composite holographic associative recall model.

item. Because cued recall depends principally on associative information, we expected output encoding to influence the correlation between recognition and recall for the local- but not for the global-match models.

Method. Methods generally followed those use in Simulation 1. During study, each item was probabilistically encoded with $p \sim \mathcal{N}(\mu_p, 0)$ and then stored according to the equations given in Table 1. We did not introduce variability in p for this simulation. The values of μ_p, N, and the resemblance criterion for recognition and for recall were set at the same levels as those in Simulation 1. The probe items themselves were perfectly encoded for cuing memory (i.e., $p = 1.0$), but were not perfectly stored.

For both local-match models we assumed that recognizing an item as old leads to the storage of the retrieved information (both members of the pair), using p_{oe}. In contrast, judging an item as "new" results in either (a) no output encoding or (b) storage of the probe item itself, also using p_{oe}. In the case of the local-match matrix model, we stored $(\tilde{\mathbf{f}} + \tilde{\mathbf{g}})(\tilde{\mathbf{f}} + \tilde{\mathbf{g}})'$ if an item was recognized and either nothing or $\tilde{\mathbf{g}}\tilde{\mathbf{g}}'$ if the item was not recognized (the tilde symbol designates the probabilistically encoded version of each item).

The global-match models do not use associative information in making item recognition judgments. Therefore, recognizing an item as old leads to the storage of the probe item itself using p_{oe}. In contrast, judging an item as new results in either (a) no output encoding or (b) storage of the probe item itself.

Results. As shown in Figure 4, output encoding of recognized items increases the observed correlation for the local-match convolution model and to a much lesser degree for the local-match matrix model. Output encoding has no effect on the global-match models. The reason for the increased correlation in the local-match models is that successful recognition of an item is based on the retrieval of the associates of that item (viz., both words in the pair). The output encoding of this associative information strengthens the same information used in cued recall, leading to an increase in $\mathcal{Q}$.

Summary of Simulations 1 and 2

The correlation between successive memory tasks does not merely reflect the degree to which those tasks tap the same information, structures, or processes in memory. Variability in the encoding of study pairs increases the correlation between successive recognition and recall (Figure 3). Output encoding increases the recognition–recall correlation for the local-match models because these models assume that an associative retrieval process underlies item recognition. There is no effect of output encoding on the global-match models because the strengthening of item information has a negligible effect on the associative information tested in recall (Figure 4).

The global-match models, which do not use associative information in item recognition, both predict near independence of recognition and recall for all but the shortest list lengths (see Figure 2). The local-match models both use associative information in item recognition. In the case of the matrix model, this leads to a high correlation between item recognition and cued recall. In the convolution models, however, the largest terms contributing to the variance in the recognition and recall processes increase with list length, while the covariance term does not. This leads the models to predict near independence of recognition and recall for lists of more than 20 pairs.

In summary, these simulations demonstrate that with substantial variability in goodness of encoding, all of the models can produce high correlations between recognition and recall. Similarly, output encoding can increase the observed correlations in the local-match models.

Empirical Tests of Encoding Variability

Whereas the foregoing analyses suggest that variability in goodness of encoding can significantly alter the predicted correlation between recognition and recall, the question of whether this factor can actually play a role in experimental studies remains unknown. Variability in goodness of encoding is always present in list learning experiments, but its magnitude can be altered. We therefore designed two experiments that manipulated the variability in goodness of encoding in an effort to determine whether, and to what extent, this variability influences the correlation between successive recognition and recall. Of critical interest was whether generalized versions of our four models that incorporated variability in goodness of encoding, variability in response criteria, output encoding, and interstimulus similarity could fit the data from these two experiments.

We manipulated goodness of encoding by varying either presentation rate (Experiment 1) or number of spaced repetitions (Experiment 2) in a mixed-list/pure-list design. Word pairs presented for longer durations or more repetitions are designated *strong,* whereas those presented for shorter durations or fewer repetitions are designated *weak.* In Experiment 1, strong pairs were presented for 8 s and weak pairs were presented for 2 s. In Experiment 2, strong pairs were presented four times and weak pairs were presented only once. In the mixed-list (high variability) condition, subjects studied lists consisting of strong and weak pairs randomly intermixed. Two pure-list conditions served as controls. In these lists, all word pairs were either strong or weak. According to the variability hypothesis, the recognition–recall correlation should be higher in the mixed list than in either of the pure lists. However, this prediction must hold in the high variability condition assuming that the strength manipulations are effective. Our interest, therefore, is in the magnitude of this effect, not its presence. In addition to testing the effect of variability on the correlation between recognition and recall, Experiment 1 also measured the magnitude of output encoding effects (Humphreys & Bowyer, 1980; Shimamura, 1985). Half of the study items did not appear on the item-recognition test, but were included in the cued-recall test. The extent that cued recall is higher for items present in the recognition test reflects some effect of Test 1 on the responses given in Test 2. As discussed earlier, the presence of such output encoding may induce further dependency between successive memory tasks.

Experiment 1

Method

Subjects. Thirty-four undergraduate students who were both English speakers and touch typists participated for payment.

Procedure. The experiment consisted of 10 study-test lists, with testing split over two sessions. The first study-test list was a practice list while the remaining 9 lists were divided into three replications of the three list types (pure weak, mixed, and pure strong). Lists were composed of words randomly selected without replacement from the Toronto Word Pool

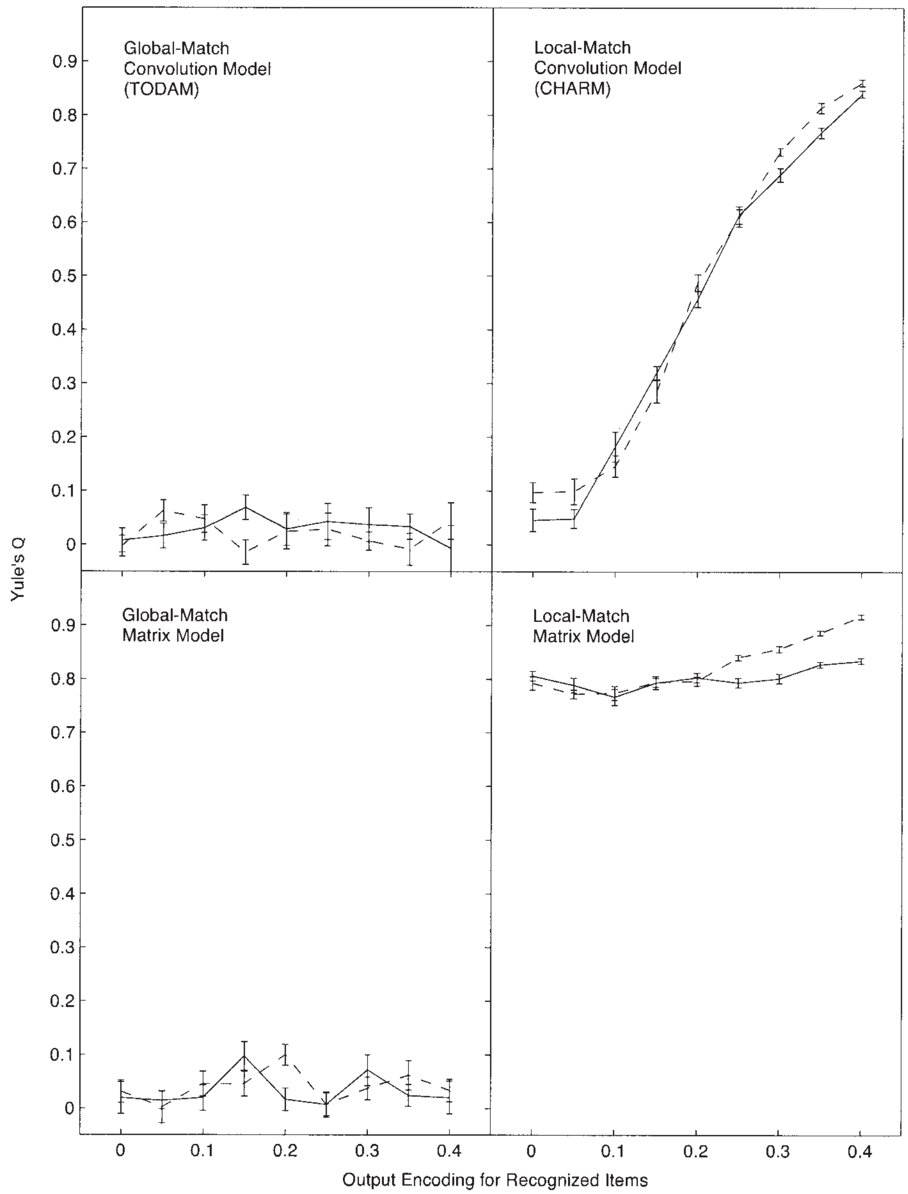

Figure 4. Effect of output encoding on the dependency between item recognition and cued recall. Each panel shows results for a different distributed memory model. Q is plotted as a function of output encoding (p_{oe}) under two different assumptions. Solid lines indicate equal output encoding for both recognized and nonrecognized items; dotted lines indicate encoding for recognized items only. Error bars represent standard error. TODAM = theory of distributed associative memory; CHARM = composite holographic associative recall model.

(Friendly, Franklin, Hoffman, & Rubin, 1982). The word pool and order of trials were randomized at the start of the first session. None of the words presented in Session 1 appeared in Session 2.

Each list consisted of a study phase in which 40 word pairs (designated A–B) were presented for varying durations. In the pure-weak lists, all pairs were shown for 2 s; in the pure-strong lists, all pairs were shown for 8 s; and in the mixed lists, half were shown for 2 s and half for 8 s. The final 4 word pairs in the study list served as a recency buffer; they were not tested later in the experiment.

A yes–no recognition test immediately followed presentation of the study list. Of the 36 study pairs tested for recall, only 18 were tested for recognition (in the mixed list, half of these items were from strong pairs and the other half were from weak pairs). Old items, chosen equally often as either the A or the B member of these 18 pairs, were intermixed with 18 distractor items. Subjects were instructed to press the *y* key for old items and the *n* key for new items.

After completing the recognition test, subjects were given a cued-recall test on all of the word pairs in the study list. If the B member of a pair was tested in recognition, recall was tested in the forward direction (i.e., A–?); if the A member of a pair was tested in recognition, recall was tested in the backward direction (i.e., B–?). In this way, recall of a word was always compared with recognition of the same word. For those pairs not present in the recognition test, half were tested in the forward and half in the backward direction. Subjects were given a maximum of 20 s to type their responses on a computer keyboard (pressing *Enter* advanced to the next probe).

Results

As shown in Table 3, list composition had a significant effect on recall and recognition performance. Lengthening the presentation rate increased recall probability and hit rates while lowering false-alarm rates. Furthermore, mixed-list performance fell between that of the pure-strong and the pure-weak lists. A comparison of recall rates for items that were and were not tested in the earlier recognition test revealed a significant advantage for recalling tested items. This result replicates the findings of Humphreys and Bowyer (1980). Supporting these observations, an analysis of variance (ANOVA) on recall accuracy revealed a significant main effect of list type, $F(2, 68) = 43.59$, $MSE = 0.014$, $p < .001$, and of prior recognition test, $F(1, 34) = 48.90$, $MSE = 0.004$, $p < .001$, but no significant interaction ($F < 1$). The d' also varied significantly with presentation rate, $F(2, 68) = 3.77$, $MSE = 0.178$, $p = .028$).

Although all three conditions exhibited moderate dependence between recognition and cued recall (Table 3), list composition did have a statistically significant effect on Q, $F(2, 68) = 3.41$, $MSE = 0.079$, $p < .05$. Tukey's honestly significant difference (HSD) tests revealed that the mixed condition produced the highest dependencies, Q (mixed) $> Q$ (pure weak), $p < .02$; Q (mixed) $> Q$ (pure strong), $p < .05$. The difference between Q in the two pure lists did not approach significance.

One may ask whether mixing strong and weak items within a list exaggerates or attenuates the effect of item strength. In associative recall and recognition tasks, such mixing of strong and weak items results in better memory for strong items and worse memory for weaker items, when compared with pure-strong and pure-weak lists (Ratcliff, Clark, & Shiffrin, 1990; Tulving & Hastie, 1972; Verde & Rotello, 2004). However, this so-called list-strength effect (LSE) is notably absent in item-recognition tasks (Murdock & Kahana, 1993; Ratcliff et al., 1990), except under special conditions (Diana & Reder, in press; Norman, 2002). Although it was not central to the goals of our study, we examined the LSE in our item-recognition and cued-recall data. To measure the LSE, we computed the ratio of the performance measure for strong to weak items in a mixed list by the same ratio in the pure lists. If this ratio of ratios (ROR) is significantly greater than 1.0, this indicates a positive LSE. As expected, there was no LSE in recognition (ROR = 1.12, $p > .10$). There was, however, a substantial LSE in cued recall (ROR = 1.42, $p < .05$). It should be noted, however, that our experiment did not specifically control for the mean study-test lag of strong and weak items in mixed versus pure lists. Not controlling for this factor could increase the magnitude of the LSE because weak items in mixed lists would be tested at a longer average lag than weak items in pure lists (Murnane & Shiffrin, 1991).

Experiment 2

In Experiment 1, variability in presentation rate produced a modest but significant increase in the recognition–recall dependency. Experiment 2 examined the effect of variable repetitions on the recognition–recall relation. In this study, pairs designated as weak appeared just once; whereas pairs designated as strong appeared four times in a spaced fashion. We also switched to a single-session, between-subject design. Several other refinements in our procedures are noted in the *Method* section.

Method

Subjects. One hundred seventy-five undergraduate students, all English speakers, participated for course credit. We assigned 42 subjects to each of the pure-strong and mixed conditions, and 91 subjects to the pure-weak condition. This increased sample size in the pure-weak condition was used because the low levels of recall led to large standard errors in the determination of Test 1–Test 2 contingencies.

Procedure. Over the course of a 1-hr session, each subject studied and was tested on three different lists. Each list consisted of 32 word pairs, randomly selected from the Toronto Word Pool. In the pure-weak lists, each pair was presented once; in the pure-strong lists, each pair was presented four times; and in the mixed list, half of the pairs were presented once and half of the pairs were presented four times. Presentation order was randomized subject to the constraint that each repeated pair was separated by at least two different pairs.

A 1-min arithmetic distractor task immediately followed the presentation of the study list. Subjects viewed equations of the form $A + B + C = D$. They pressed *y* if the equation was correct and *n* if it was incorrect (for half of the equations, the value of D was off by two). The experimenter stressed accuracy in the instructions, but subjects had to respond to each equation within 2 s. After each trial, a computer-generated tone (high pitch for correct, low pitch for incorrect) provided feedback.

Following the distractor task, subjects made yes–no recognition judgments on the 32 B items intermixed with an equal number of lures. The

Table 3
Descriptive Statistics for Experiments 1 and 2

Condition	$P(C_t)$	HR	FAR	d'	Q	$P(C_u)$
Experiment 1						
Pure weak	.26 (.03)	.76 (.02)	.17 (.02)	1.84 (0.11)	.49 (.05)	.20 (0.02)
Mixed	.35 (.03)	.77 (.01)	.15 (.02)	1.86 (0.09)	.66 (.05)	.29 (0.03)
Pure strong	.45 (.04)	.81 (.02)	.14 (.02)	2.09 (0.13)	.54 (.04)	.39 (0.03)
Experiment 2						
Pure weak	.26 (.01)	.70 (.01)	.18 (.01)	1.60 (0.05)	.52 (.02)	
Mixed	.42 (.03)	.75 (.02)	.11 (.01)	2.01 (0.08)	.63 (.03)	
Pure strong	.54 (.04)	.83 (.02)	.08 (.01)	2.80 (0.17)	.49 (.04)	

Note. Numbers in parentheses are standard errors of the mean. Yule's Q was determined separately for each subject. $P(C_t)$ = probability of cued recall for items that were tested in the recognition phase; HR = hit rate; FAR = false-alarm rate; $P(C_u)$ = probability of cued recall for untested items.

cued-recall task was administered immediately following this recognition task. For each A–B pair, the A item was displayed and subjects were instructed to speak the B item[8] into a microphone. If subjects did not remember the target item, they were to say "pass." The computer digitally recorded subjects' responses for later scoring of response accuracy and latency.

Results

As shown in Table 3, both recognition and recall performance increased with repetition. An ANOVA revealed a significant main effect of list type on recall probability, $F(2, 172) = 44.50$, $MSE = 0.027$, $p < .001$ (Tukey's HSD tests confirmed that each of the differences among conditions was highly significant, $p < .01$). Repetition also influenced d', $F(2, 172) = 42.80$, $MSE = 0.48$, $p < .001$, with HSD tests confirming significant differences on all of the comparisons ($p < .01$). As in Experiment 1, we examined the LSE in both our item-recognition and cued-recall data. Once again, we found no significant LSE in item recognition (ROR = 0.88, $p > .10$), but a significant positive LSE in cued recall (ROR = 1.34, $p < .05$).

As in Experiment 1, we expected to find the highest recognition–recall dependencies in the mixed-list condition. As shown in Table 3, Q differed significantly across the three list types, $F(2, 172) = 5.10$, $MSE = 0.045$, $p < .01$. Tukey's HSD tests revealed that Q in the mixed condition was significantly higher than in either the pure-weak ($p = .02$) or pure-strong conditions ($p < .01$). As is clear from the means, there was almost no difference in Q between pure-strong and pure-weak lists ($p > .5$). Despite the powerful repetition manipulation used in this study, the absolute magnitude of the increase in dependency in the mixed condition was quite modest, as it was in Experiment 1.

Model Fits

As shown in the two reported experiments, increasing variability in either number of presentations or presentation rate produced a reliable increase in the correlation between successive recognition and recall tasks. Whereas the pure lists replicated the classic ≈ .5 correlation between item recognition and cued recall, the correlation increased to ≈ .6 under conditions of experimentally induced encoding variability. Also, pairs that were not tested in the recognition phase (a manipulation in Experiment 1) were less likely to be recalled than those pairs that were tested.

The base versions of our four DMMs cannot even account for the classic moderate correlation between recognition and subsequent recall. They either over- or underpredict the recognition–recall correlation. Only the local-match matrix model predicts a strong positive intertask correlation, but this predicted correlation is far too high. The other three DMMs predict near independence between recognition and recall. However, our formal analysis of intertask correlations revealed that a number of factors not considered in the base versions of the models can strongly modulate the recognition–recall correlation. In particular, variability in goodness of encoding can increase the correlation in all four models, and output encoding can increase the correlation in the local-match models. Variability in the retrieval process can decrease the correlation between successive tasks because the retrieval events represent independent sources of variation. We therefore asked whether generalized versions of the four DMMs can fit the data from Experiments 1 and 2.

Method

Each of the four models was simultaneously fit[9] to the full-contingency tables in the pure-weak, pure-strong, and mixed conditions as these aspects of the data were central to our theoretical analysis.[10] For each parameter set, the study and test of 36 pairs of list items (32 for Experiment 2) were simulated 200 times for each condition. In both experiments, we simulated strength using the probabilistic encoding mechanism common to all four models. For each strong word pair, the proportion of elements stored, p, was drawn from a truncated normal distribution with mean μ_{p_s} and variance $\sigma^2_{p_s}$; for weak pairs, the mean and variance parameters were μ_{p_w} and $\sigma^2_{p_w}$. These parameters were the same regardless of list composition (mixed vs. pure). Note that this empirical approach to modeling strength does not address the complicated question of how repetitions affect learning (see Rizzuto & Kahana, 2001, for a discussion of modeling learning in DMMs).

We varied 10 parameters to optimize the goodness of fit in each of the four models: mean and standard deviation of probabilistic encoding for strong lists (μ_{p_s}, σ_{p_s}) and for weak lists (μ_{p_w}, σ_{p_w}), output encoding for recognized items (p_{oe}), interitem similarity, and the mean and standard deviation of the resemblance criteria for recognition and for recall.

Results

Table 4 shows each of the model's fit to data from Experiment 1. Each of the four models can produce a moderate correlation between recognition and recall (≈ .5) in pure lists while simultaneously providing good fits to the hit rate and the overall level of recall. Each of the models also predicts a somewhat higher correlation in the mixed condition, as found in the experiment. Only the local-match models, however, could account for the higher recall of items tested in the recognition phase. This effect was predicted because output encoding stores associative information during the recognition phase for the local-match models, but only stores item-specific information, which does not help much with recall, for the global-match models.

[8] To keep these procedures consistent with those used in previous successive testing experiments, we tested memory for the B member of each *A–B* pair both in recognition and in recall.

[9] An evolutionary algorithm (Mitchell, 1996) with an initial population of 256 points (that were uniformly distributed in the parameter space) evolved until the best fitness (smallest value of $\sum_i [(\text{observed}_i - \text{expected}_i)/\sigma_{\text{observed}_i})]^2$) did not change from one generation to the next. At the end of each generation, 10% of the top parameter vectors were saved, 40% were copies of the top 10% with Gaussian point mutations, 30% were recombinations of the top 10%, and the remaining 20% were randomly generated parameter vectors. The models converged to a reliable parameter set after approximately 10 generations.

[10] We did not fit other aspects of the data, such as the recognition false-alarm rate or the intrusion rate in cued recall. The values of N needed to achieve the observed levels of recall would necessarily predict d' values that are far higher than those observed in the data, especially in the global-match models. This problem may be solved by adding a temporal context term to the storage equation (e.g., Dennis & Humphreys, 2001; Howard & Kahana, 2002; Murdock, 1997) and implementing a continuous model in which the memory vector/matrix is not reset at the start of each list (Murdock & Kahana, 1993). Such extensions might be fruitful directions for future work.

Table 4
Observed and Predicted Values for Each of the Four Models Fit to Experiment 1

	Pure strong				Pure weak				Mixed			
Model	$P(C_t)$	$P(C_u)$	HR	Q	$P(C_t)$	$P(C_u)$	HR	Q	$P(C_t)$	$P(C_u)$	HR	Q
Observed	.45	.39	.81	.54	.26	.20	.76	.49	.35	.29	.77	.66
GM conv.	.39	.38	.79	.53	.21	.21	.76	.60	.30	.30	.78	.62
LM conv.	.45	.40	.79	.49	.26	.20	.76	.56	.37	.32	.77	.70
GM matrix	.43	.43	.79	.46	.20	.19	.78	.45	.32	.31	.76	.61
LM matrix	.43	.36	.80	.57	.27	.21	.76	.55	.34	.29	.77	.63

Note. We simultaneously fit each of the models to probability of cued recall for tested, $P(C_t)$, and for untested, $P(C_u)$, items, hit rate (HR), and Yule's Q for each of the following list types: pure strong, pure weak, and mixed. GM = global match; conv. = convolution; LM = local match.

Table 5 gives the best fitting parameter values for each of the four models. To fit these data, all of the models required some level of variability in goodness of encoding, as reflected in the σ_p-parameter values for strong and weak pairs. This factor, which will increase the correlation between recognition and recall, is essential for the two global-match models that predict near independence of recognition and recall for the simplified model derivations shown in Figure 1. The local-match models both assume that recognized items are weakly stored, as indicated by the small positive values of parameter p_{oe}. This parameter enables these models to account for the enhanced recall of items that were tested during the recognition phase.[11] The potency of output encoding for the local-match models is approximately half of the potency of a single presentation during study (e.g., a weak pair). Larger values of this parameter would have predicted a larger effect of the recognition test on subsequent recall than observed in the data.

Several other parameters turned out to be important in fitting these data. Variability in the recognition and recall criteria served to decrease the correlation between recognition and recall. This factor was used heavily by the local-match matrix model that predicts a very high correlation between item recognition and cued recall in the derivations shown in Figure 1. Interitem similarity was called into play by both of the local-match models. This factor acts to increase the correlation between recognition and recall by increasing the number of overlapping terms in the covariance.

Table 6 shows each of the model's fit to data from Experiment 2. As with Experiment 1, each of the four models accounted for the moderate correlation between recognition and recall ($\approx$ .5) in the pure-list conditions while simultaneously providing good fits to the hit rate and the overall level of recall. Each of the models also predicted a higher correlation in the mixed condition, as found in the experiment.

Table 7 gives the best fitting parameter values for each of the four DMMs. The overall fit, while somewhat worse than the fit to Experiment 1, was nonetheless reasonably good. The best fitting parameter values obtained in these fits was in fairly good agreement with those from the fits to Experiment 1. The global-match models assume a significant degree of variability in goodness of encoding to account for the moderate correlation between recognition and recall. As in the previous simulation, both of the local-match models make significant use of output encoding as reflected in the p_{oe} parameter. All of the models assume some degree of variability in the recognition and recall criteria. Variability in the recognition criterion was higher for the two local-match models than for the global-match models. Interitem similarity was called into play by both of the local-match models and also by the global-match convolution model.

Summary of Model Fits

Whereas base versions of the four DMMs cannot reproduce the moderate observed correlation between recognition and recall, generalized versions that take account of variability during storage and retrieval, output encoding, and interitem similarity were all able to qualitatively capture the complex pattern of experimental results. Each model did so by relying more or less heavily on parameters that tend to increase or decrease the predicted correlation between tasks. For example, had we turned off the variability in the recognition and recall decisions, the local-match matrix model would have predicted a significantly higher correlation between recognition and recall. Similarly, without variability in goodness of encoding, the global-match matrix model would have predicted a significantly lower correlation between recognition and recall. Although the architecture of the models determines a base level of correlation, as derived in Table 2, more realistic models of recognition and recall should include other factors that can modulate this correlation. As illustrated in Figure 2, any factor that increases common variance will tend to raise the correlation, and any factor that increases unique variance will tend to decrease the correlation. Here we see the manifestation of this statistical truism in the detailed execution of the four DMMs.

Whereas the two reported experiments manipulated variability in goodness of encoding and output encoding, we did not manipulate other factors, such as retrieval variability, that would likely have placed tighter constraints on the four models. Indeed, such data might have enabled us to demonstrate that some (or all) of the models cannot fit the pattern of intertask correlations. This remains an open target for future research.

[11] As would be expected, the small value of p_{oe} has no effect on cued recall for this model. It may be that nonzero value is an artifact of the optimization method because the likelihood surface is very insensitive to the value of this parameter.

Table 5
Best Fitting Parameter Values for Each of the Four Models Fit to Experiment 1

Model	Pair sim.	Strong μ_p	Strong σ_p	Weak μ_p	Weak σ_p	p_{oe}	Recall criteria	Recall sigma	Recognition criteria	Recognition sigma
GM conv.	.00	.57	.16	.44	.17	.07	.39	.20	−1.10	.20
LM conv.	.27	.53	.08	.33	.08	.16	.23	.20	−0.82	.16
GM matrix	.00	.77	.20	.50	.13	.00	.31	.04	−0.75	.11
LM matrix	.10	.63	.10	.48	.08	.23	.24	.35	−1.20	.32

Note. sim. = similarity; GM = global match; conv. = convolution; LM = local match.

General Discussion

Recognition and recall serve as the two standard measures of episodic memory. Although both tasks measure intentional retrieval of previously experienced events, they could not be more different. Numerous experimental manipulations differentially affect recognition and recall. Examples of these dissociations include the word frequency effect (Gregg, 1976; Kinsbourne & George, 1974; MacLeod & Kampe, 1996), list strength effect (Ratcliff et al., 1990), intentional encoding effects (Glenberg & Bradley, 1979; Schwartz & Humphreys, 1974), associative interference (Dyne, Humphreys, Bain, & Pike, 1990), context effects (Godden & Baddeley, 1975, 1980), subject age (Craik & McDowd, 1987), and damage to the medial-temporal lobe (Hirst, 1986; Vargha-Khadem et al., 1997). These dissociations, in turn, have led theorists to advocate for a distinction between item-specific and relational information (Humphreys, 1978; Murdock, 1974), with familiarity-based retrieval of item-specific information and recollection of relational information (Yonelinas, 1997; Yonelinas, Kroll, Dobbins, Lazzara, & Knight, 1998).

Analyzing the Recognition–Recall Relation

Our approach has been to use computational models of memory to help understand the relation between recognition and recall. We selected four DMMs for detailed analyses. The four models differ along two critical dimensions: the mechanism of association and the processes underlying recognition. We refer to the four models as (a) global-match convolution, (b) local-match convolution, (c) global-match matrix, and (d) local-match matrix. The convolution models use the mathematical operation of convolution to form associations; the matrix models use matrix operations (Hebbian learning) to store and retrieve associations. The global-match models represent dual-process models with completely separate recognition and recall mechanisms. The local-match models assume that the same machinery underlies both recognition and recall. The global-match convolution model is a version of Murdock's (1982) model. The global-match matrix model is a version of the Humphreys, Bain, and Pike (1989) matrix model (without context).[12] The local-match convolution model is a version of Metcalfe's (1985) model. Finally, the local-match matrix model is a variant of the models presented in Rizzuto and Kahana (2001) and Kahana (2002). Each of the chosen models represents items as vectors of features, and assumes that item and associative information is stored in a common distributed memory system. Each model uses probabilistic encoding to model the effects of repetition on learning.

Although these types of DMMs have been used extensively to account for recognition and recall data (Humphreys, Bain, & Pike, 1989; Metcalfe, 1985; Murdock, 1982), the basic assumptions about the dependence or independence of the information supporting these tasks had not been carefully evaluated. We thus began our analysis by analytically deriving the correlation between recognition and recall in each of the models. This exercise revealed an interesting pattern, with the local-match matrix model predicting a high level of dependency, and the other models predicting a moderate level of dependency for short lists, but tending toward independence for lists consisting of 20 or more pairs of items. The moderate dependency predicted for short lists was a consequence of the assumption that memory was set to zero at the start of learning. With a more reasonable continuous memory assumption (e.g., Murdock & Kahana, 1993), one would expect these models to predict low correlations even for shorter lists.

To test these assumptions about the relation between recognition and recall, we turned to an extensive body of literature on successive recognition and recall tests. In these studies, the correlation between recognition and recall at the level of individual A–B pairs is measured by testing recognition of the B items in a first test phase, and then giving a subsequent cued-recall test, with each A item serving as the cue for its mate. Tabulating subjects' Test 1 and Test 2 responses in a 2 × 2 contingency table reveals a moderate level of dependency between item recognition and cued recall. That is, whether an item was recognized on the first test is moderately predictive of whether it will be recalled in the second test. Quantified using a measure such as Yule's Q, the correlation between recognition and recall is about .5.[13]

Interpreting the Recognition–Recall Relation

The mapping between theory and data is problematized by factors that can influence the correlation between any two mea-

[12] We did not implement list context in the four DMMs because only some of these models have been formulated to include a representation of contextual features. If one were to extend the implementation of context in Murdock (1997) or Humphreys, Bain, and Pike (1989) to the other models, the result would be an increase in interitem similarity or featural overlap between all list items. In simulations of the effect of similarity on the recognition–recall correlation, we found that interitem similarity increased the predicted correlation between recognition and recall for the local-match models, but decreased the correlation for the global-match models.

[13] If the recall test is given first, then the B items that are recalled would have received an extra opportunity for encoding, thus boosting later recognition performance and the correlation between tasks.

Table 6
Observed and Predicted Values for Each of the Four Models Fit to Experiment 2

	Pure strong			Pure weak			Mixed		
Model	$P(C)$	HR	Q	$P(C)$	HR	Q	$P(C)$	HR	Q
Observed	.54	.83	.49	.26	.71	.53	.42	.75	.63
GM conv.	.62	.81	.58	.26	.71	.47	.46	.76	.64
LM conv.	.44	.80	.48	.25	.74	.53	.39	.76	.73
GM matrix	.48	.74	.38	.28	.72	.59	.38	.73	.60
LM matrix	.48	.78	.58	.26	.73	.53	.38	.76	.62

Note. We simultaneously fit each of the models to probability of cued recall, $P(C)$, hit rate (HR), and Yule's Q for each of the following list types: pure strong, pure weak, and mixed. GM = global match; conv. = convolution; LM = local match.

surements. Correlations between successive tests are produced by variability. Common sources of variance increase the correlation, and separate sources of variance decrease the correlation. In the successive testing paradigm, encoding conditions will be common to the storage of information supporting both recognition and recall, and may thus boost the correlation between tasks. With recognition and recall phases of the task being widely separated in time, retrieval conditions can introduce unique variance and thus lower the correlation. Furthermore, output encoding will raise the correlation between successive tests under certain conditions.

We considered such factors in modeling recognition and recall data from successive tests. For all four models considered in this article, variability in goodness of encoding predicted an increase in the correlation between recognition and recall. Output encoding also predicted an increase in the correlation for the local-match but not for the global-match models. This is because only the local-match models assume that recognition involves a retrieval and restorage of the associations as well as the items. Variability in the encoding of the test probes, or the decision process, decreased the correlation between recognition and recall in all of the models.

Experiments 1 and 2 confirmed that at least one of these factors, variability in goodness of encoding, can be manipulated to increase the correlation between recognition and recall, as predicted for all four models. The effect was modest, increasing the recognition–recall dependency by about 20% (from $Q \approx .5$ to $Q \approx .6$). Although future research will be needed to pin down the effects of other variables on the recognition–recall relation, these initial results led us to examine whether the four models, when given the flexibility to vary these factors, could mimic the observed pattern of results.

In fitting the data from Experiments 1 and 2, all four classes of models could predict the moderate level of dependency observed experimentally and its modest increase when strong and weak items were mixed in a given list. The local-match models were also able to explain the increase in recall associated with testing an item during recognition. Thus, both those models that assume independence of the underlying information supporting recognition and recall, and those that assume a very high correlation, can account for the moderate correlations observed experimentally.

Contingency Analyses and the Classification of Memory

Researchers have used the method of successive tests, and contingency analyses, to investigate problems in human memory other than the recognition–recall relation. In what may be considered the earliest example of this work, Estes (1960) demonstrated an extremely high correlation between the outcomes of successive learning trials. Such an effect was viewed as consistent with the view that individual pairs of items were learned in a single trial, and that once such pairs were learned they were seldom forgotten. Later, researchers used contingency analyses to examine the source of associative interference in the Barnes and Underwood (1959) "unlearning" paradigm. In these studies, subjects learn a list of A–B pairs to a performance criterion. Then they study a list of A–C pairs, in which the previously studied A items are each paired with a new item. (The degree of A–C learning is an experimental parameter in these studies.) Finally, subjects are probed with each A item to recall both B and C in any order. According to the Melton–Underwood *unlearning-recovery hypothesis,* the decrease in B recall following A–C learning results from specific unlearning

Table 7
Best Fitting Parameter Values for Each of the Four Models Fit to Experiment 2

		Strong		Weak						
Model	Pair sim.	μ_p	σ_p	μ_p	σ_p	p_{oe}	Recall criteria	Recall sigma	Recognition criteria	Recognition sigma
GM conv.	.19	.68	.19	.41	.12	.00	.19	.19	−1.00	.14
LM conv.	.33	.62	.00	.39	.00	.21	.10	.19	−0.87	.20
GM matrix	.01	.81	.19	.57	.20	.00	.29	.19	−0.56	.12
LM matrix	.15	.61	.00	.44	.00	.08	.21	.15	−1.00	.21

Note. sim. = similarity; GM = global match; conv. = convolution; LM = local match.

of the individual A–B associations, thus predicting a negative contingency between recall of B and C items. Dapolito (1967) found that contrary to these predictions, recall of B and C is nearly independent across a wide range of experiments (see Kahana, 2000, for a review). These findings played a key role in the demise of the classic associative unlearning theory of forgetting.

More recently, Kahana and colleagues (Kahana, 2002; Rizzuto & Kahana, 2001) used the successive testing method to test the independent associations hypothesis (Wolford, 1971)—the view that forward and backward associations are formed independently. This view predicts that the correlation between forward recall and subsequent backward recall of the same pair should be significantly lower than the correlation between pairs tested in the same direction. Contingency analyses revealed that contrary to this view, forward and backward recall was almost perfectly correlated and was as highly correlated as pairs tested in the same direction. This finding was taken to support the associative symmetry hypothesis (Asch & Ebenholtz, 1962; Kahana, 2002), which sees associations as a holistic conjunction of both A and B items. Caplan, Glaholt, and McIntosh (2005) replicated Kahana's finding of symmetric recall of pairs, but showed that for recall of triples and serial lists the correlation between forward and backward probes was significantly reduced.

Finally, contingency analyses have been applied widely in the study of the relation between implicit and explicit memory. Whereas explicit memory tasks, such as recognition and recall, directly probe subjects' memory for a studied event, implicit tasks measure memory without reference to a particular study episode. As reviewed previously, performance on successive explicit memory tasks is moderately to highly correlated. In contrast, when one task is implicit (e.g., fragment completion) and the other is explicit (e.g., recognition), the correlations are often near zero (Hayman & Tulving, 1989a; Tulving & Hayman, 1995; Tulving et al., 1982, but see also Ostergaard, 1992). This independence is also found when a fragment completion task is administered twice, first with implicit memory instructions (e.g., no reference to the study episode) and then later with explicit memory instructions (Hayman & Tulving, 1989a, 1989b). Such findings of independence between implicit and explicit memory tests have been taken as strong support for a memory systems view in which explicit tasks probe information in an episodic memory system and implicit tasks access information in a functionally and anatomically independent perceptual representation system (Tulving & Schacter, 1991).

The analysis presented here suggests an alternative account of these findings. In the fragment completion task, which is the source of much data on implicit memory, subjects can complete the fragment without reference to information stored in the original study episode. Variability in completion is thus a function of independent variation in item difficulty for different fragments or nonoverlapping fragments of the same word. Without focusing retrieval on a studied episode, the relative contribution of test variability will be greater than that of variability in storage (which produces priming). As shown by our modeling of the recognition–recall relation, test variability lowers the correlation between tasks. As demonstrated by Experiments 1 and 2, and by our modeling work, variability in the goodness of encoding raises the correlation between tasks. Consequently, the factors that could produce correlations between explicit tasks (viz., variability in the goodness of encoding) will have little impact on implicit tasks (which depend largely on the difficulty of completion for a given fragment). According to this account, there is need for only one memory system; the information that produces priming comes from the same memory system that enables explicit recognition and recall. The only time when significant dependencies would be observed between implicit and explicit tasks is when the size of the priming effect is large relative to the overall fragment completion rates. Even then, the obtained dependency would be less than one would find in successive explicit memory tasks.

The successive testing paradigm is complex and unforgiving. The correlation between information tested in successive tasks is only one of several factors that contribute to the observed intertask dependency. If one can experimentally estimate the values of encoding and test variability, as well as the effects of output encoding, then dependencies between successive tasks will provide important constraints on memory models. The present work represents a first step in this direction.

The models examined in this article are in many ways too simple. More sophisticated models that incorporate temporal or positional context (e.g., Brown, Preece, & Hulme, 2000; Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, & Usher, 2005; Dennis & Humphreys, 2001; Howard & Kahana, 2002; Murdock, 1997), and/or that allow for dual-process mechanisms in recognition (e.g., Norman & O'Reilly, 2003; Reder et al., 2000; Yonelinas, 1996), may make different predictions concerning recognition–recall dependencies. As our analyses illustrate, application of such models to the recognition–recall relation, or the relation between any pair of memory tasks, will first require explicit assumptions about variability and output encoding—factors that have been largely ignored in the memory modeling literature. Furthermore, additional data on the effects of these factors will be needed before data on intertask correlations can help us to accept or reject specific models or modeling assumptions.

References

Anderson, J. A. (1970). Two models for memory organization using interacting traces. *Mathematical Biosciences, 8,* 137–160.

Anderson, J. A. (1973). A theory for the recognition of items from short memorized lists. *Psychological Review, 80,* 417–438.

Anderson, J. A., Silverstein, J. W., Ritz, S. A., & Jones, R. S. (1977). Distinctive features, categorical perception, and probability learning: Some applications of a neural model. *Psychological Review, 84,* 413–451.

Asch, S. E., & Ebenholtz, S. M. (1962). The principle of associative symmetry. *Proceedings of the American Philosophical Society, 106,* 135–163.

Bahrick, H. P. (1970). Two-phase model for prompted recall. *Psychological Review, 77,* 215–222.

Barnes, J. M., & Underwood, B. J. (1959). Fate of first-list associations in transfer theory. *Journal of Experimental Psychology, 58,* 97–105.

Batchelder, W. H., & Riefer, D. M. (1995). A multinomial modeling analysis of the recognition-failure paradigm. *Memory & Cognition, 23,* 611–630.

Begg, I. (1979). Trace loss and the recognition failure of unrecalled words. *Memory & Cognition, 7,* 113–123.

Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review, 107,* 127–181.

Caplan, J. B., Glaholt, M., & McIntosh, A. R. (2005). *Linking associative and list memory: Pairs versus triples.* Manuscript submitted for publication.

Chappell, M., & Humphreys, M. (1994). An autoassociative neural network for sparse representations: Analysis and application to models of recognition and cued recall. *Psychological Review, 101,* 103–128.

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin and Review, 3,* 37–60.

Craik, F. I. M., & McDowd, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 474–479.

Dapolito, F. J. (1967). Proactive effects with independent retrieval of competing responses. *Dissertation Abstracts International, 27,* 2522–2523.

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

Dennis, S., & Humphreys, M. (2001). A context noise model of episodic word recognition. *Psychological Review, 108,* 452–478.

Diana, R. A., & Reder, L. M. (in press). The list strength effect: A contextual competition account. *Memory & Cognition.*

Donnelly, R. E. (1988). Priming effects in successive episodic tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 256–265.

Dyne, A. M., Humphreys, M. S., Bain, J. D., & Pike, R. (1990). Associative interference effects in recognition and recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 813–824.

Estes, W. K. (1960). Learning theory and the new "mental chemistry." *Psychological Review, 67,* 207–223.

Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review, 9,* 59–85.

Flexser, A. J., & Tulving, E. (1978). Retrieval independence in recognition and recall. *Psychological Review, 85,* 153–171.

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.

Glenberg, A. M., & Bradley, M. M. (1979). Mental contiguity. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 88–97.

Godden, D., & Baddeley, A. (1975). Context-dependent memory in two natural environments: On land and under water. *British Journal of Psychology, 66,* 325–331.

Godden, D., & Baddeley, A. (1980). When does context influence recognition memory? *British Journal of Psychology, 71,* 99–104.

Gregg, V. (1976). Word frequency, recognition and recall. In J. Brown (Ed.), *Recall and recognition* (pp. 183–216). Oxford, England: Wiley.

Hayman, C. G., & Tulving, E. (1989a). Contingent dissociation between recognition and fragment completion: The method of triangulation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 228–240.

Hayman, C. G., & Tulving, E. (1989b). Is priming in fragment completion based on a "traceless" memory system? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 941–956.

Hintzman, D. (1981). Simpson's paradox and the analysis of memory retrieval. *Psychological Review, 87,* 398–410.

Hintzman, D. (1987). Recognition and recall in MINERVA 2: Analysis of the "recognition-failure" paradigm. In P. Morris (Ed.), *Modelling cognition* (pp. 215–229). New York: Wiley.

Hintzman, D. (1988). Judgments of frequency and recognition memory in multiple-trace memory model. *Psychological Review, 95,* 528–551.

Hintzman, D. (1992). Mathematical constraints on the Tulving–Wiseman function. *Psychological Review, 99,* 536–542.

Hintzman, D., & Hartry, A. L. (1990). Item effects in recognition and fragment completion: Contingency relations vary for different subsets of words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 965–969.

Hirst, W. (1986). Recognition and recall in amnesics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 445–451.

Hockley, W. E., & Cristi, C. (1996). Tests of encoding tradeoffs between item and associative information. *Memory & Cognition, 24,* 202–216.

Hockley, W. E., & Murdock, B. B. (1987). A decision model for accuracy and response latency in recognition memory. *Psychological Review, 94,* 341–358.

Holdstock, J., Mayes, A., Roberts, N., Cezayirli, E., Isaac, C., O'Reilly, R., & Norman, K. (2002). Under what conditions is recognition spared relative to recall after selective hippocampal damage in humans? *Hippocampus, 12,* 341–351.

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

Humphreys, M. S. (1978). Item and relational information: A case for context independent retrieval. *Journal of Verbal Learning and Verbal Behavior, 17,* 175–187.

Humphreys, M. S., Bain, J. D., & Pike, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic, and procedural tasks. *Psychological Review, 96,* 208–233.

Humphreys, M. S., & Bowyer, P. A. (1980). Sequential testing effects and the relation between recognition and recognition failure. *Memory & Cognition, 8,* 271–277.

Humphreys, M. S., Pike, R., Bain, J. D., & Tehan, G. (1989). Global matching: A comparison of the SAM, Minerva II, Matrix, and TODAM models. *Journal of Mathematical Psychology, 33,* 36–67.

Jordan, M. I. (1986). An introduction to linear algebra in parallel distributed processing. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Foundations* (Vol. 1, pp. 365–422). Cambridge, MA: MIT Press.

Kahana, M. J. (2000). Contingency analyses of memory. In E. Tulving & F. I. M. Craik (Eds.), *Oxford handbook of human memory* (pp. 323–384). New York: Oxford Press.

Kahana, M. J. (2002). Associative symmetry and memory theory. *Memory & Cognition, 30,* 823–840.

Kinsbourne, M., & George, J. (1974). The mechanism of the word-frequency effect on recognition memory. *Journal of Verbal Learning and Verbal Behavior, 13,* 63–69.

Lisman, J., Jensen, O., & Kahana, M. J. (2001). Toward a physiologic explanation of behavioral data on human memory. In C. Hölscher (Ed.), *Neuronal mechanisms of memory formation* (pp. 195–223). Cambridge, England: Cambridge University Press.

MacLeod, C. M., & Kampe, K. (1996). Word frequency effects on recall, recognition, and word fragment completion tests. *Journal of Experimental Psychology, 22,* 132–142.

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

Metcalfe, J. (1985). Levels of processing, encoding specificity, elaboration, and CHARM. *Psychological Review, 92,* 1–38.

Metcalfe, J. (1991). Recognition failure and the composite memory trace in CHARM. *Psychological Review, 98,* 529–553.

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

Murdock, B. B. (1974). *Human memory: Theory and data.* Hillsdale, NJ: Erlbaum.

Murdock, B. B. (1979). Convolution and correlation in perception and memory. In L. G. Nilsson (Ed.), *Perspectives in memory research: Essays in honor of Uppsala University's 500th anniversary* (pp. 105–119). Hillsdale, NJ: Erlbaum.

Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review, 89,* 609–626.

Murdock, B. B. (1992). Item and associative information in a distributed memory model. *Journal of Mathematical Psychology, 36,* 68–99.

Murdock, B. B. (1995). Similarity in a distributed memory model. *Journal of Mathematical Psychology, 39,* 251–264.

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

Murdock, B. B., & Kahana, M. J. (1993). Analysis of the list strength effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 689–697.

Murdock, B. B., & Lamon, M. (1988). The replacement effect: Repeating some items while replacing others. *Memory & Cognition, 16,* 91–101.

Murnane, K., & Shiffrin, R. M. (1991). Interference and the representation of events in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 855–874.

Nilsson, L. G., & Gardiner, J. M. (1991). Memory theory and the boundary conditions of the Tulving–Wiseman law. In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock* (pp. 57–74). Hillsdale, NJ: Erlbaum.

Nilsson, L. G., & Gardiner, J. M. (1993). Identifying exceptions in a database of recognition failure studies from 1973 to 1992. *Memory & Cognition, 21,* 397–410.

Nilsson, L. G., Law, J., & Tulving, E. (1988). Recognition failure of recallable unique names: Evidence for an empirical law of memory and learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 266–277.

Norman, K. A. (2002). Differential effects of list strength on recollection and familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 1083–1094.

Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary learning systems approach. *Psychological Review, 110,* 611–646.

Ostergaard, A. L. (1992). A method for judging measures of stochastic dependence: Further comments on the current controversy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 413–420.

Pike, R. (1984). Comparison of convolution and matrix distributed memory systems for associative recall and recognition. *Psychological Review, 91,* 281–294.

Ratcliff, R., Clark, S. E., & Shiffrin, R. M. (1990). List-strength effect: I. Data and discussion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 163–178.

Reder, L. M., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, R., & Hiraki, K. A. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember-know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 294–320.

Rizzuto, D. S., & Kahana, M. J. (2001). An autoassociative neural network model of paired associate learning. *Neural Computation, 13,* 2075–2092.

Schwartz, R. M., & Humphreys, M. S. (1974). Recognition and recall as a function of instructional manipulations of organization. *Journal of Experimental Psychology, 102,* 517–519.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM—Retrieving effectively from memory. *Psychonomic Bulletin and Review, 4,* 145.

Shimamura, A. P. (1985). Problems with the finding of stochastic independence as evidence for multiple memory systems. *Bulletin of the Psychonomic Society, 23,* 506–508.

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.

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

Tulving, E., & Hastie, R. (1972). Inhibition effects of intralist repetition in free recall. *Journal of Experimental Psychology, 92,* 297–304.

Tulving, E., & Hayman, C. G. (1995). On measurement of priming: What is the correct baseline? *European Journal of Cognitive Psychology, 7,* 13–18.

Tulving, E., & Schacter, D. L. (1990, Jan 19). Priming and human memory systems. *Science, 247,* 301–305.

Tulving, E., Schacter, D. L., & Stark, H. A. (1982). Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 336–342.

Tulving, E., & Thompson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352–373.

Tulving, E., & Wiseman, S. (1975). Relation between recognition and recognition failure of recallable words. *Bulletin of the Psychonomic Society, 6,* 79–82.

Vargha-Khadem, F., Gadian, D. G., Watkins, K. E., Connely, A., Van Paesschen, W., & Mishkin, M. (1997, July 18). Differential effects of early hippocampal pathology on episodic and semantic memory. *Science, 277,* 376–380.

Verde, M. F., & Rotello, C. M. (2004). Strong memories obscure weak memories in associative recognition. *Psychonomic Bulletin & Review, 11,* 1062–1066.

Weber, E. U. (1988). Expectation and variance of item resemblance distributions in a convolution–correlation model of distributed memory. *Journal of Mathematical Psychology, 32,* 1–43.

Weber, E. U., & Murdock, B. B. (1989). Priming in a distributed memory system: Implications for models of implicit memory. In S. Lewandowsky, J. C. Dunn, & K. Kirsner (Eds.), *Implicit memory: Theoretical issues* (pp. 87–89). Hillsdale, NJ: Erlbaum.

Wiseman, S., & Tulving, E. (1976). Encoding specificity: Relation between recall superiority and recognition failure. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 349–361.

Wolford, G. (1971). Function of distinct associations for paired-associate performance. *Psychological Review, 78,* 303–313.

Yonelinas, A. P. (1996). Dissociating recollection and familiarity in recognition memory. *Dissertation Abstracts International, 57,* 2193.

Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition, 25,* 747–763.

Yonelinas, A. P., Kroll, N. E. A., Dobbins, I. G., Lazzara, M., & Knight, R. T. (1998). Recollection and familiarity deficits in amnesia: Convergence of remember-know, process dissociation, and receiver operating characteristic data. *Neuropsychology, 12,* 323–339.

(*Appendixes follow*)

Appendix A

Solving for the Correlation Between Recognition and Recall

This appendix derives expressions for the theoretical correlation between item recognition and cued recall in the global- and local-match variants of the convolution–correlation and matrix memory models. These derivations rely on component level variance and covariance expressions presented in Appendix B.

Global-Match Convolution–Correlation Model

Following Murdock (1982), the memory vector sums the item and heteroassociative information for all of the word pairs in the list. Assuming, as is common, that memory is reset at the start of each list yields the following storage equation for a list of L word pairs:

$$\mathbf{m} = \sum_{i=1}^{L} (\mathbf{f}_i + \mathbf{g}_i + \mathbf{f}_i * \mathbf{g}_i).$$

Cued recall is achieved by probing the memory vector with either $\mathbf{f}$ or $\mathbf{g}$ for retrieval of its mate.

$$\mathbf{f}_j\#\mathbf{m} = \mathbf{f}_j\#\mathbf{f}_j + \mathbf{f}_j\#\mathbf{g}_j + \mathbf{f}_j\#\mathbf{f}_j * \mathbf{g}_j + \sum_{i \neq j}^{L} (\mathbf{f}_j\#\mathbf{f}_i + \mathbf{f}_j\#\mathbf{g}_i + \mathbf{f}_j\#\mathbf{f}_i * \mathbf{g}_i).$$

The correlation of an item with itself (e.g., $\mathbf{f}_j\#\mathbf{f}_j$) yields the vector $\delta = (\ldots, x, 1, x, \ldots)$ with $E(x) = 0$ and $V(x) = 1/N$. For large N, $\delta \approx (\ldots, 0, 1, 0, \ldots)$. Similarly, the correlation of two different random vectors (e.g., $\mathbf{f}_j\#\mathbf{g}_j$) is a vector whose elements all have an expected value of zero. Consequently, for large N, $\mathbf{f}_j\#\mathbf{m} \approx \mathbf{f}_j\#\mathbf{f}_j * \mathbf{g}_j = \delta * \mathbf{g}_j = \mathbf{g}_j$. Thus, regardless of the number of associations stored in the memory vector, an approximate representation of the target item can be recovered. To retrieve the desired item, the noisy retrieved information must be cleaned up. Automatic deblurring is not a feature of the linear DMMs considered here. Nonetheless, one can compute recall probabilities by comparing the retrieved information with a lexicon of possible target items. The most similar item (as measured by the dot product) that falls within a region around the expected value of the target item is chosen as the retrieved item.

Recognition decisions are based on the resemblance of a probe item with the memory vector. For instance, if $\mathbf{g}$ is a probe item, the dot product $\mathbf{g} \cdot \mathbf{m}$ provides a measure of the "strength" of item $\mathbf{g}$. This strength can serve as input to a decision system for recognition judgments (e.g., Hockley & Murdock, 1987). Although the expected dot product of vectors representing items ($\mathbf{f}$ or $\mathbf{g}$), heteroassociations ($\mathbf{f} * \mathbf{g}$), and autoassociations ($\mathbf{f} * \mathbf{f}$) is zero (i.e., $E[(\mathbf{f} * \mathbf{f}) \cdot (\mathbf{f} * \mathbf{g})] = E[\mathbf{f} \cdot (\mathbf{f} * \mathbf{g})] = E[\mathbf{f} \cdot (\mathbf{f} * \mathbf{f})] = 0$), this does not imply independence of recognition and recall (contrary to Murdock, 1982).

The correlation between recall and recognition is given by

$$\rho = \frac{\mathrm{cov}[\mathbf{g} \cdot \mathbf{m}, (\mathbf{f}\#\mathbf{m}) \cdot \mathbf{g}]}{\sqrt{\mathrm{var}[\mathbf{g} \cdot \mathbf{m}]\mathrm{var}[(\mathbf{f}\#\mathbf{m}) \cdot \mathbf{g}]}} = \frac{\mathrm{cov}[\mathbf{g} \cdot \mathbf{m}, (\mathbf{f} * \mathbf{g}) \cdot \mathbf{m}]}{\sqrt{\mathrm{var}[\mathbf{g} \cdot \mathbf{m}]\mathrm{var}[(\mathbf{f} * \mathbf{g}) \cdot \mathbf{m}]}}.$$

This simplification is achieved by using the recognition–recall identity for convolution models (Murdock, 1992). To determine this correlation, variance and covariance expressions are derived. To compute the total covariance, we break it down into components. Those components that result in nonzero covariances are shown in Table A1.

The values for $A1$ and $A2$ are derived in Appendix B. The total covariance is given by $\mathrm{cov}[(\mathbf{f}_j * \mathbf{g}_j) \cdot \mathbf{m}, \mathbf{g}_j \cdot \mathbf{m}] = 3.75N^{-1} + N^{-2} + 0.25N^{-3}$. The variance of the recognition term and recall terms, calculated using expressions derived in Weber (1988), are given in Table 2.

In this and subsequent models, interitem similarity is modeled by generating random vectors that have a correlation, ρ, to a hidden prototype vector. If $\mathbf{z}$ denotes the prototype vector and exemplars $\mathbf{f} = \rho\mathbf{z} + \sqrt{1 - \rho^2}\mathbf{u}$, and $\mathbf{g} = \rho\mathbf{z} + \sqrt{1 - \rho^2}\mathbf{v}$, then $E[\mathbf{f} \cdot \mathbf{g}] = \rho^2$ (for details, see Murdock, 1995). In several of the simulations reported in the body of the text, interitem similarity was modeled according to these equations. The derivations presented in this appendix do not consider interitem similarity.

Local-Match Convolution–Correlation Model

Following Metcalfe (1985), the memory vector sums the autoassociative and heteroassociative information for all of the word pairs in the list. The storage equation is given by

$$\mathbf{m} = \sum_{i=1}^{L} (\mathbf{f}_i * \mathbf{f}_i + \mathbf{g}_i * \mathbf{g}_i + 2\mathbf{f}_i * \mathbf{g}_i) = \sum_{i=1}^{L} (\mathbf{f}_i + \mathbf{g}_i)^{*2}.$$

Cued recall is achieved by probing the memory vector with either $\mathbf{f}$ or $\mathbf{g}$ for retrieval of its mate: $\mathbf{f}_j\#\mathbf{m} = \mathbf{f}_j + 2\mathbf{g}_j +$ noise. Because autoassociative information is stored in memory, the retrieved information consists of both the probe and the target items. Response probability is proportional to the resemblance of the retrieved information to the target item: $(\mathbf{f}\#\mathbf{m}) \cdot \mathbf{g} = (\mathbf{f} * \mathbf{g}) \cdot \mathbf{m}$.

Recognition decisions are based on a two-stage process: First, the probe is correlated with the memory vector to retrieve the associated information, $\mathbf{r} = \mathbf{g}_j\#\mathbf{m} = 2\mathbf{f}_j + \mathbf{g}_j +$ noise. Next, the retrieved information is matched against the probe item, $\mathbf{r} \cdot \mathbf{g}_j$. If its resemblance to the retrieved information exceeds some criterion, the probe is recognized.

This recognition process is referred to as a *local-match process* because only the recovered information enters the memory comparison. In contrast, a *global-match process,* compares the probe item with all of the items stored in the memory trace.

Table A1
Covariance Terms for the Global-Match Convolution Model

Term	$\mathbf{f} \cdot (\mathbf{f} * \mathbf{g})$	$\mathbf{g} \cdot (\mathbf{f} * \mathbf{g})$	$(\mathbf{f} * \mathbf{g}) \cdot (\mathbf{f} * \mathbf{g})$	$\mathbf{u} \cdot (\mathbf{f} * \mathbf{g})$	$(\mathbf{f} * \mathbf{g}) \cdot (\mathbf{u} * \mathbf{u})$
$\mathbf{f} \cdot \mathbf{f}$	0	0	$A1$	0	0
$\mathbf{f} \cdot \mathbf{g}$	0	0	0	0	0
$\mathbf{f} \cdot (\mathbf{f} * \mathbf{g})$	$A2$	0	0	0	0
$\mathbf{f} \cdot \mathbf{u}$	0	0	0	0	0
$\mathbf{f} \cdot (\mathbf{u} * \mathbf{v})$	0	0	0	0	0

Note. Nonzero terms are derived in Appendix B.

Table A2
Covariance Terms for the Local-Match Convolution Model

Term	$(\mathbf{f} * \mathbf{g}) \cdot (\mathbf{f} * \mathbf{f})$	$(\mathbf{f} * \mathbf{g}) \cdot (\mathbf{g} * \mathbf{g})$	$(\mathbf{f} * \mathbf{g}) \cdot (\mathbf{f} * \mathbf{g})$	$(\mathbf{f} * \mathbf{g}) \cdot (\mathbf{u} * \mathbf{v})$	$(\mathbf{f} * \mathbf{g}) \cdot (\mathbf{u} * \mathbf{v})$
$(\mathbf{f} * \mathbf{f}) \cdot (\mathbf{f} * \mathbf{f})$	0	0	*B*1	0	0
$(\mathbf{f} * \mathbf{f}) \cdot (\mathbf{g} * \mathbf{g})$	0	0	*B*2	0	0
$(\mathbf{f} * \mathbf{f}) \cdot (\mathbf{f} * \mathbf{g})$	*B*3	*B*4	0	*B*5	0
$(\mathbf{f} * \mathbf{f}) \cdot (\mathbf{u} * \mathbf{u})$	0	0	*B*6	0	0
$(\mathbf{f} * \mathbf{f}) \cdot (\mathbf{u} * \mathbf{v})$	0	0	0	0	0

Note. Nonzero terms are derived in Appendix B.

Using the recognition–recall identity (Murdock, 1992), the correlation coefficient between item recognition and cued recall is given by

$$\rho = \frac{\mathrm{cov}[(\mathbf{f}_j * \mathbf{f}_j) \cdot \mathbf{m}, (\mathbf{f}_j * \mathbf{g}_j) \cdot \mathbf{m}]}{\sqrt{\mathrm{var}[(\mathbf{f}_j * \mathbf{f}_j) \cdot \mathbf{m}]\mathrm{var}[(\mathbf{f}_j * \mathbf{g}_j) \cdot \mathbf{m}]}}. \tag{A1}$$

The total covariance is obtained by summing all pairwise covariances. Each nonzero pairwise covariance is indicated in Table A2.

The values for *B*1–*B*6 are derived in Appendix B. The variance terms, calculated using expressions derived in Weber (1988), are given in Table 2. Substituting the variance and covariance expressions into Equation A1 then gives the correlation between item recognition and cued recall.

Global-Match Matrix Model

The matrix model of Humphreys, Bain, and Pike (1989) extends the formalism proposed by Anderson (1970) to a full-fledged model of item and associative recognition, cued recall, fragment cued recall, and a variety of other tasks. The matrix model is structurally similar to Murdock's (1982) global-match convolution model. The major difference is that the Humphreys et al. model uses context-to-item associations to model the distinction between episodic and semantic memory. To facilitate comparison across the four models, we implement a version of the matrix model that does not incorporate context. We therefore write the storage equation as

$$W = \sum_{i=1}^{L} (\mathbf{f}_i\mathbf{g}_i' + \mathbf{g}_i\mathbf{f}_i' + \mathbf{f}_i\mathbf{r}_i' + \mathbf{r}_i\mathbf{g}_i').$$

Probability of successfully recognizing item $\mathbf{g}$ is proportional to the match of $\mathbf{g}$ with the memory matrix; symbolically, $(\mathbf{r}\mathbf{g}_p') \cdot W$, where the dot notation is used to take a dot circle product of matrices.

For cued recall, the probe is multiplied with the memory matrix to get an approximate representation of the target item. The match of the retrieved item with the memory information $(W\mathbf{f}_p) \cdot \mathbf{g}_p'$ then determines the probability of successful cued recall. Expanding at the component level reveals an associative recognition–cued recall identity for the matrix model:

$$W\mathbf{f}_p \cdot \mathbf{g}_p' = \sum_{k=1}^{N}\left(\sum_{j=1}^{N}\sum_{i=1}^{L}(f_j^i g_k^i + g_j^i f_k^i + f_j^i + g_k^i)f_k^p\right)g_j^p$$

$$= \sum_{k=1}^{N}\sum_{j=1}^{N} W_{jk} f_k^p g_j^p = \mathbf{g}_p\mathbf{f}_p' \cdot W.$$

Consequently, the correlation between item recognition and cued recall can be written as follows:

$$\rho = \frac{\mathrm{cov}(W\mathbf{f}_p \cdot \mathbf{g}_p', W \cdot \mathbf{r}\mathbf{g}_p')}{\sqrt{\mathrm{var}(W\mathbf{f}_p \cdot \mathbf{g}_p']\mathrm{var}(W \cdot \mathbf{r}\mathbf{g}_p')}} = \frac{\mathrm{cov}(W \cdot \mathbf{g}_p\mathbf{f}_p', W \cdot \mathbf{r}\mathbf{g}_p')}{\sqrt{\mathrm{var}(W \cdot \mathbf{g}_p\mathbf{f}_p']\mathrm{var}(W \cdot \mathbf{r}\mathbf{g}_p')}}.$$

Local-Match Matrix Model

Matrix products can be used to support either autoassociation or heteroassociation. Following Rizzuto and Kahana (2001), we consider a matrix model that uses the autoassociation of a sum of items to store both autoassociative and heteroassociative information. This shares the property of associative symmetry (Kahana, 2002; Rizzuto & Kahana, 2001) with the convolution models. That is, the *F*–*G* association will be just as strong as the *G*–*F* association. The storage equation for this model is given by

$$W = \sum_{i=1}^{L} (\mathbf{f}_i + \mathbf{g}_i)(\mathbf{f}_i + \mathbf{g}_i)'.$$

As in the local-match convolution model, recognition decisions are based on a two-stage process: First, the cue item retrieves the stored trace. Second, this retrieved information is matched against the cue $W\mathbf{g}_p \cdot \mathbf{g}_p$. If the item was successfully stored, the retrieved information should be similar to the probe item. The information supporting item recognition performance is proportional to

$$\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{i=1}^{L}(f_j^i g_k^i g_j^p g_k^p + g_j^i f_k^i g_j^p g_k^p + f_j^i f_k^i g_j^p g_k^p + g_j^i g_k^i g_j^p g_k^p)$$

Cued recall works exactly as it does in the global-match matrix model—the probe is multiplied with the memory matrix to get an approximate representation of the target item. The match of the retrieved item with the memory information then determines the probability of successful cued recall. The information supporting cued-recall performance is proportional to

$$\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{i=1}^{L}(f_j^i g_k^i f_j^p g_k^p + g_j^i f_k^i f_j^p g_k^p + f_j^i f_k^i f_j^p g_k^p + g_j^i g_k^i f_j^p g_k^p).$$

The variance and covariance terms for the local-match matrix model are derived in Appendix B.

(*Appendixes continue*)

Appendix B

Variance and Covariance Derivations

This appendix derives variance and covariance components of item and associative information supporting recognition and recall. $W, X, Y,$ and $Z \sim \mathcal{N}(0, 1/N)$ denote independent and identically distributed random variables. Boldfaced symbols **f, g, u,** and **v** denote vectors, L denotes list length (in pairs), **m** is a memory vector, W is a memory/weight matrix, and $\mathbf{f}'$ denotes the transpose of **f**.

Using the moment-generating function, $M_x(\theta) = e^{(1/2)\sigma^2\theta^2}$, one can derive the expectations for any power of a Gaussian random variable. Expanding this function as a Taylor series and then differentiating at zero yields the following: $E(X^{2n+1}) = 0$ and $E(X^{2n}) = \sigma^{2n}(2n)!/2^n n!$, $n \in Z$. Expectations of even powers of X are given by $E(X^2) = \sigma^2$, $E(X^4) = 3\sigma^4$, $E(X^6) = 15\sigma^6$; for odd powers of X the expectation is zero.

In counting terms, the total number of terms in $(\mathbf{f} * \mathbf{g}) \cdot (\mathbf{u} * \mathbf{v})$ is given by $(2N^3 + 4N^2 + N)/3$. For the means and variances of the convolution–correlation models, component terms have already been derived by Weber (1988) and are not given here.

$$\mathrm{cov}[(\mathbf{f} * \mathbf{g}) \cdot (\mathbf{f} * \mathbf{g}),\ \mathbf{f} \cdot \mathbf{f}] = N^2\mathrm{cov}(X^2Y^2, X^2) = \frac{2}{N}$$

$$\begin{aligned}\mathrm{cov}[(\mathbf{f} * \mathbf{g}) \cdot \mathbf{f},\ (\mathbf{f} * \mathbf{g}) \cdot \mathbf{f}] &= N\mathrm{cov}(X^2Y, X^2Y) \\ &\quad + (N^2 - N)\mathrm{cov}(X^2Y, Z^2Y) \\ &\quad + (0.75N^2 - N)\mathrm{cov}(XYZ, XYZ) \\ &= \frac{7N^2 + 4N + 1}{4N^3}\end{aligned}$$

$$\begin{aligned}\mathrm{cov}[(\mathbf{f} * \mathbf{f}) \cdot (\mathbf{f} * \mathbf{f}),\ (\mathbf{f} * \mathbf{g}) \cdot (\mathbf{f} * \mathbf{g})] &= N^2\mathrm{cov}(X^4, X^2Y^2) \\ &\quad + 2N(N^2 - N)2\mathrm{cov}(X^2Y^2, X^2Z^2) = 8N^{-1} + 4N^{-2}\end{aligned}$$

$$\begin{aligned}\mathrm{cov}[(\mathbf{f} * \mathbf{f}) \cdot (\mathbf{g} * \mathbf{g}),\ (\mathbf{f} * \mathbf{g}) \cdot (\mathbf{f} * \mathbf{g})] &= N\mathrm{cov}[X^2Y^2, X^2Y^2] \\ + 2(N^2 - N)\mathrm{cov}(X^2Y^2, X^2Z^2) &+ 4(N^2 - N)\mathrm{cov}(XYZW, XYZW) = 8N^{-2}\end{aligned}$$

$$\begin{aligned}&\mathrm{cov}[(\mathbf{f} * \mathbf{g}) \cdot (\mathbf{f} * \mathbf{f}),\ (\mathbf{f} * \mathbf{g}) \cdot (\mathbf{f} * \mathbf{f})] \\ &= N\mathrm{cov}(X^3Y, X^3Y] + 4(N^2 - N)\mathrm{cov}(X^3Y, Z^2XY) \\ &\quad + \left(4(N^2 - N) + \frac{(N-1)^2}{2}\right)\mathrm{cov}(X^2YZ, X^2YZ) \\ &\quad + 4(N^2 - N)(N - 2)\mathrm{cov}(X^2YZ, W^2YZ) \\ &\quad + \frac{4N^3 - 7N^2 + 14N - 3}{12}\mathrm{cov}(WXYZ, WXYZ) \\ &= \frac{32N^3 + 67N^2 - 92N + 3}{6N^4}\end{aligned}$$

$$\begin{aligned}&\mathrm{cov}[(\mathbf{f} * \mathbf{g}) \cdot (\mathbf{f} * \mathbf{f}),\ (\mathbf{f} * \mathbf{g}) \cdot (\mathbf{g} * \mathbf{g})] \\ &= N\mathrm{cov}(X^3Y, Y^3X) + 4(N^2 - N)\mathrm{cov}(X^3Y, XYZ^2) \\ &\quad + 4(N^2 - N)(N - 1)\mathrm{cov}(X^2YZ, W^2YZ) \\ &\quad + \frac{(N-1)^2}{2}\mathrm{cov}(X^2YZ, W^2YZ) = \frac{8N^3 + 9N^2 + 1}{2N^4}\end{aligned}$$

$$\begin{aligned}&\mathrm{cov}[(\mathbf{f} * \mathbf{g}) \cdot (\mathbf{u} * \mathbf{u}),\ (\mathbf{f} * \mathbf{g}) \cdot (\mathbf{f} * \mathbf{f})] \\ &= N\mathrm{cov}(X^3Y, W^2XY) + (N^2 - N)2\mathrm{cov}(X^2YZ, W^2YZ) \\ &\quad + \frac{(N-1)^2}{2}\mathrm{cov}(X^2YZ, W^2YZ) = \frac{5N^2 + 1}{2N^4}\end{aligned}$$

$$\begin{aligned}&\mathrm{cov}[(\mathbf{f} * \mathbf{f}) \cdot (\mathbf{u} * \mathbf{u}),\ (\mathbf{f} * \mathbf{g}) \cdot (\mathbf{f} * \mathbf{g})] \\ &= N\mathrm{cov}[X^2Y^2, X^2Z^2] + (N^2 - N)\mathrm{cov}(X^2Y^2, X^2Z^2) = 2N^{-2}\end{aligned}$$

$$\begin{aligned}&\mathrm{cov}[W \cdot (\mathbf{g}_p\mathbf{f}'_p),\ W \cdot (\mathbf{r}\mathbf{g}'_p)] \\ &= \mathrm{cov}\left(\sum_{k=1}^{N}\sum_{j=1}^{N} W_{jk}g_j^p f_k^p,\ \sum_{k=1}^{N}\sum_{j=1}^{N} W_{jk}g_k^p\right) \\ &= \mathrm{cov}\left[\begin{array}{l}\sum_{k=1}^{N}\sum_{j=1}^{N}\sum_{i=1}^{L}(f_j^i g_k^i + g_j^i f_k^i + f_j^i + g_k^i)g_j^p f_k^p, \\ \sum_{k'=1}^{N}\sum_{j'=1}^{N}\sum_{i'=1}^{L}(f_{j'}^{i'} g_{k'}^{i'} + g_{j'}^{i'} f_{k'}^{i'} + f_{j'}^{i'} + g_{k'}^{i'})g_{k'}^p\end{array}\right] \\ &= 2 + 4N^{-1} + 4N^{-2}\end{aligned}$$

$$\begin{aligned}&\mathrm{var}[W \cdot (\mathbf{g}'_p\mathbf{f}'_p)] \\ &= \mathrm{var}\left[\sum_{k=1}^{N}\sum_{j=1}^{N}\sum_{i=1}^{L}(f_j^i g_k^i + g_j^i f_k^i + f_j^i + g_k^i)g_j^p f_k^p,\right] \\ &= \mathrm{cov}\left[\begin{array}{l}\sum_{k=1}^{N}\sum_{j=1}^{N}\sum_{i=1}^{L}(f_j^i g_k^i + g_j^i f_k^i + f_j^i + g_k^i)g_j^p f_k^p, \\ \sum_{k'=1}^{N}\sum_{j'=1}^{N}\sum_{i'=1}^{L}(f_{j'}^{i'} g_{k'}^{i'} + g_{j'}^{i'} f_{k'}^{i'} + f_{j'}^{i'} + g_{k'}^{i'})g_{j'}^{p'} f_{k'}^{p'}\end{array}\right] \\ &= (2L + 6)N^{-1} + (2L + 18)N^{-2} + (2L + 6)N^{-3}\end{aligned}$$

$$\begin{aligned}\mathrm{var}(W \cdot \mathbf{r}\mathbf{g}'_p) &= \mathrm{var}\left[\sum_{k=1}^{N}\sum_{j=1}^{N}\sum_{i=1}^{L}(f_j^i g_k^i + g_j^i f_k^i + f_j^i + g_k^i)g_k^p\right] \\ &= \mathrm{cov}\left[\begin{array}{l}\sum_{k=1}^{N}\sum_{j=1}^{N}\sum_{i=1}^{L}(f_j^i g_k^i + g_j^i f_k^i + f_j^i + g_k^i)g_k^p, \\ \sum_{k'=1}^{N}\sum_{j'=1}^{N}\sum_{i'=1}^{L}(f_{j'}^{i'} g_{k'}^{i'} + g_{j'}^{i'} f_{k'}^{i'} + f_{j'}^{i'} + g_{k'}^{i'})g_{k'}^p\end{array}\right] \\ &= (L + 2)N + (L + 2) + (2L + 5)N^{-1} + (2L + 6)N^{-2}\end{aligned}$$

$$\begin{aligned}&\mathrm{cov}[(W\mathbf{f}_p) \cdot \mathbf{g}_p,\ (W\mathbf{g}_p) \cdot \mathbf{g}_p] \\ &= \mathrm{cov}\left(\begin{array}{l}\sum_{k=1}^{N}\sum_{j=1}^{N}\sum_{i=1}^{L} f_j^i g_k^i f_k^p g_j^p + g_j^i f_k^i f_k^p g_j^p + f_j^i f_k^i f_k^p g_j^p + g_j^i g_k^i f_k^p g_j^p, \\ \sum_{k'=1}^{N}\sum_{j'=1}^{N}\sum_{i'=1}^{L} f_{j'}^{i'} g_{k'}^{i'} g_{k'}^p g_{j'}^p + g_{j'}^{i'} f_{k'}^{i'} g_{k'}^p g_{j'}^p + f_{j'}^{i'} f_{k'}^{i'} g_{k'}^p g_{j'}^p + g_{j'}^{i'} g_{k'}^{i'} g_{k'}^p g_{j'}^p\end{array}\right) \\ &= 8N^{-1} + (8L + 30)N^{-2} + (12L + 30)N^{-3}\end{aligned}$$

$$\mathrm{var}[(W\mathbf{f}_p)\cdot\mathbf{g}_p] = \mathrm{cov}\left(\begin{array}{l} \displaystyle\sum_{k=1}^{N}\sum_{j=1}^{N}\sum_{i=1}^{L} f_j^i g_k^i f_k^p g_j^p + g_j^i f_k^i f_k^p g_j^p + f_j^i f_k^i f_k^p g_j^p + g_j^i g_k^i f_k^p g_j^p, \\ \displaystyle\sum_{k=1}^{N}\sum_{j=1}^{N}\sum_{i=1}^{L} f_{j'}^{i'} g_{k'}^{i'} f_{k'}^p g_{j'}^p + g_{j'}^{i'} f_{k'}^{i'} f_{k'}^p g_{j'}^p + f_{j'}^{i'} f_{k'}^{i'} f_{k'}^p g_{j'}^p + g_{j'}^{i'} g_{k'}^{i'} f_{k'}^p g_{j'}^p \end{array}\right)$$

$$= 8N^{-1} + (12L + 22)N^{-2} + (4L^2 + 12L + 22)N^{-3}$$

$$\mathrm{var}[(W\mathbf{g}_p)\cdot\mathbf{g}_p] = \mathrm{cov}\left(\begin{array}{l} \displaystyle\sum_{k=1}^{N}\sum_{j=1}^{N}\sum_{i=1}^{L} f_j^i g_k^i g_k^p g_j^p + g_j^i f_k^i g_k^p g_j^p + f_j^i f_k^i g_k^p g_j^p + g_j^i g_k^i g_k^p g_j^p, \\ \displaystyle\sum_{k'=1}^{N}\sum_{j'=1}^{N}\sum_{i'=1}^{L} f_{j'}^{i'} g_{k'}^{i'} g_{k'}^p g_{j'}^p + g_{j'}^{i'} f_{k'}^{i'} g_{k'}^p g_{j'}^p + f_{j'}^{i'} f_{k'}^{i'} g_{k'}^p g_{j'}^p + g_{j'}^{i'} g_{k'}^{i'} g_{k'}^p g_{j'}^p \end{array}\right)$$

$$= 12N^{-1} + (24L + 50)N^{-2} + (8L^2 + 40 + 54)N^{-3}.$$

Received September 26, 2003
Revision received January 6, 2005
Accepted March 3, 2005 ■