Psychonomic Bulletin & Review
1996, 3 (1), 37–60

Global matching models of recognition memory: How the models match the data

STEVEN E. CLARK
University of California, Riverside, California

and

SCOTT D. GRONLUND
University of Oklahoma, Norman, Oklahoma

We present a review of global matching models of recognition memory, describing their theoretical origins and fundamental assumptions, focusing on two defining properties: (1) recognition is based solely on familiarity due to a match of test items to memory at a global level, and (2) multiple cues are combined interactively. We evaluate the models against relevant data bearing on issues including the representation of associative information, differences in verbal and environmental context effects, list-length, list-strength, and global similarity effects, and ROC functions. Two main modifications to the models are discussed: one based on the representation of associative information, and the other based on the addition of recall-like retrieval mechanisms.

This paper reviews a class of models collectively called *global matching models*.[1] These models have been important in guiding a decade of empirical and theoretical work in human memory. In the present paper, we focus on episodic recognition memory. We trace the origins of models, discussing empirical data and theoretical problems that motivated the development of global memory models. We describe the models, both in terms of their common properties and at the level of specific assumptions. We show how these specific assumptions underlie each model's predictions, and we evaluate each model against standard results as well as current data. We conclude by discussing other current frameworks for recognition memory, possible modifications to global matching models, and future work.

OVERVIEW

Included in this class of models are Murdock's (1982, 1983) theory of distributed associative memory (TODAM), Eich's (1982) composite holographic associative recall model (CHARM),[2] Raaijmakers and Shiffrin's (1980, 1981; Gillund & Shiffrin, 1984, extended it to recognition) search of associative memory (SAM) model, Pike's (1984) Matrix model, and Hintzman's (1984, 1988) MINERVA 2. We will also describe more recent developments to these models, including Humphreys, Bain, and Pike (1989), and Shiffrin, Ratcliff, and Clark (1990).

In all of these models, the cues available at test are combined into a single joint probe of memory. In recognition, the test items are not used as cues to retrieve particular items from memory but instead are used to access memory broadly, activating information stored about multiple events in parallel. The result of this global level access to memory is a scalar index that may be construed as any of the following: (1) the familiarity of the test item, (2) the match of the test item with the contents of memory, or (3) the activation of memory produced by the test item. The label that we attach to this scalar is not important for present purposes, and we will use the terms *familiarity*, *match*, *activation*, and *strength* more or less interchangeably.

At this level, global matching models are quite simple and are clear variants of signal-detection models. The familiarity of a given class of test items will be distributed with mean μ and standard deviation σ. Performance is largely a function of the means and standard deviations of the familiarity distributions for targets (μ_T and σ_T) and distractors (μ_D and σ_D) and can be described in terms of d', where $d' = (\mu_T - \mu_D)/\sigma_D$. From this, it is clear that performance can be specified for test conditions for which the means and variances can be specified.[3]

This brief description of global matching models makes two claims: one regarding the manner in which multiple cues are combined to probe memory, and the other regarding the information that is accessed. We refer to these as the *interactive-cue* assumption and the *global matching* assumption. Before discussing these assumptions in detail, we describe the empirical domain in which these assumptions have been evaluated.[4]

This research was supported by National Science Foundation Grant DBS 9120911 to S.E.C. We wish to thank Doug Hintzman, Mike Humphreys, Ben Murdock, and Rich Schweickert for their valuable reviews of a lengthy manuscript, and Rich Shiffrin for helpful comments on an earlier version of the paper. Correspondence should be addressed to S. E. Clark, Psychology Department, University of California, Riverside, CA 92521 (e-mail: clark@ucrac1.ucr.edu), or to S. D. Gronlund, Psychology Department, University of Oklahoma, Norman, OK 73019 (e-mail: sgronlund@uoknor.edu).

Domain

These models have been applied to a wide range of experimental phenomena in recognition, recall, frequency estimation, serial order, categorization, depression, person perception, picture recognition, and so on. However, for present purposes, we will restrict our discussion of the models to recognition memory. Unlike recall, recognition testing does not require subjects to assemble retrieved information into overt responses. It is reasonable to assume that the fewer the number of intervening processes, the clearer the window to the representation of information in memory.

It is useful to establish some terms and notation at the outset that we will use throughout the paper. For example, in a pair study paradigm, subjects study a list of pairs, *AB*, *CD*, *EF*, and so on (where each letter denotes a word). *Item* recognition requires subjects to distinguish between words that were studied (*A*) and new words that were not on the study list (*X*). *Associative* recognition (sometimes called *pair* recognition) requires subjects to discriminate *intact* from *rearranged* test pairs, where *AB* is an intact pair and *AD* is a rearranged test pair consisting of items recombined from different study pairs.

Origins of the Models

Global memory models combined aspects of two earlier model types—search models and direct-access models—which had complementary successes and failures. The problem is understood in terms of two characteristics of recognition memory: (1) decisions are made quickly, and (2) recognition judgments are based not only on the characteristics of the test item but also on the characteristics of *other* items in memory.

Search models (e.g., Tulving, 1976) assume that events are stored separately in memory. At test, each memory item is retrieved and compared with the test item. The subject responds "yes" if a match is found and "no" otherwise. However, this serial comparison process is too slow for recognition. Within the framework of such a model, high-confidence accurate "no" responses would result from an exhaustive search comparing the test item with each of the items in memory to find zero matches. This model would produce very long response times, contrary to data showing that a proportion of correct rejection responses can be made very quickly, and with high confidence (Atkinson & Juola, 1974; Glucksberg & McCloskey, 1981).

The speed problem can be solved by assuming that the test item is retrieved not by a serial search, but rather is accessed directly. Such *direct-access* recognition models (e.g., Kintsch, 1970) were formulated within the framework of associative networks. In such models, concepts (words) are represented as nodes, and relations between concepts are represented as pathways between nodes. Recall operates as a search from node to node along associative pathways. For recognition, however, access to the network is cued by the test word itself, allowing direct access to the relevant node. The strength of information stored at the node is used to make a recognition decision. We will refer to this particular direct-access model as a *local match* model (in contrast to global match models) because the information relevant to the recognition of a given item is stored locally at the node that corresponds to that item. Access is direct because the slow node-to-node search is bypassed, and therefore recognition judgments can be made quickly. However, this speed comes at a cost. In its simplest form, this direct-access model cannot easily account for the effects of other items in recognizing a particular test item. As we shall see, recognition of a given item is affected by other items in memory, and by other (nontarget) items presented at the time of retrieval.

Other Items in Memory

In general, recognition of a given item, *A*, decreases as other items are added to memory (Bowles & Glanzer, 1983; Strong, 1912). This result is called a *list-length effect*. Other studies have varied the similarity of other items in memory. We refer to these as *global similarity effects*, because the similarity to the test probe is distributed across multiple list items. These studies sometimes come under the heading of prototype effects. For example, category prototypes (which were not studied) show high false-alarm rates (Bransford & Franks, 1971; Posner & Keele, 1970). In a related study, Hintzman (1988) showed that hit and false-alarm rates both increase as the number of similar items in a list is increased. These results all indicate that recognition of *A* is a function not only of the characteristics of *A* but also of the characteristics of other items in memory. It should be noted, however, that the sensitivity to other items in memory is not without boundaries, as shown by subjects' ability to focus recognition on a particular list and reject items presented in other lists as new (e.g., Hintzman & Waters, 1970; Murnane & Shiffrin, 1991a; Ratcliff, Clark, & Shiffrin, 1990).

Other Items at Retrieval

Effects of other items at retrieval are typically described as *context effects*. For a list of items studied as pairs, *AB*, *CD*, and so on, recognition of a given item, *B*, is better if it is tested with its original pair member (*A*) than if it is tested alone or with a different item (Humphreys, 1976; Light & Carter-Sobell, 1970; Thomson, 1972; Tulving & Thomson, 1971). Similarly, response times to call test item *B* "old" are faster if the presentation of *B* is preceded by pair member *A* (McKoon & Ratcliff, 1979; Neely & Durgunoglu, 1985). This kind of facilitation is termed an *episodic priming effect* and may be viewed as a reaction-time analogue of the context effects shown by Thomson and Tulving and others.

All of these results, taken together, show that recognition is due in part to other items. Local match models cannot account for this sensitivity to other items without some kind of modification. The ideal model is one that takes other items into consideration but allows recognition decisions to be made quickly. Search models are sensitive to other items in memory but are slow; local match models are insensitive to other items in memory but are able to account for fast response times.

Given this complementarity, it is not surprising that two-process models of recognition that combined search and direct-access were proposed. Context-based facilitation could be produced because subjects could use the context item A to recall the test item B, even if B could not be recognized directly (Humphreys, 1978; Mandler, 1980). Similarly, list-length effects can be easily produced within the framework of a search model, because errors increase the more there is to search through (Atkinson & Juola, 1974).

Gillund and Shiffrin (1984) empirically examined the necessity for a two-process account of recognition.[5] Presumably, if subjects are required to respond very quickly, the relative contribution of a search process should decrease, and recognition decisions should be based primarily on the direct-access component. Conversely, if subjects are forced to respond slowly, the relative contribution of search should increase.

Experimental manipulations that selectively reduce or eliminate the search component (but not the direct-access component) should show interactions with variables that are differentially sensitive to search. In their experiments, Gillund and Shiffrin (1984) manipulated list length and orienting task, among other variables, and required subjects to respond quickly (within 900 msec) or slowly (only after 1.5 sec had elapsed). Speeded recognition led to an overall decrease in d', but no interactions were observed. These results were inconsistent with a two-process model, leading Gillund and Shiffrin to develop SAM as a single-process model of recognition.

The major strength of global memory models is that they are sensitive to other items in memory, and they can in principle[6] produce fast responses. Thus, one process is able to account for results that had previously been taken as evidence for a two-factor theory. Although the global matching models possess a recall mechanism, the models assume that a single global matching process is sufficient to account for recognition.

SPECIFIC MODELS

Up to this point, we have treated the various global matching models as variants of a single model type, emphasizing their similarities rather than their differences. However, while all the models share the properties of global matching and interactive-cue combination, the manner in which these properties are instantiated varies considerably across models. Thus, while a "big picture" may emerge from viewing the models as a forest, they can be evaluated only at the level of the trees.

In this section, we will describe four global matching models. The section distinguishes between separate-storage (SAM and MINERVA 2) and distributed-storage models (TODAM and the Matrix model). Global matching in separate- and distributed-storage models is shown in Figure 1. In a separate-storage model (top panel), events $E_1, E_2, \ldots, E_N$ are stored in corresponding separate, distinct memory representations $M_1, M_2, \ldots, M_N$. The test item is matched to each stored representation, and these separate matches are summed at test to produce an index of the global familiarity of the test item. In distributed-storage models (bottom panel), multiple events are distributed over a common set of units. In TODAM and the Matrix model, multiple events are summed together at study to form a single composite memory. The test item is matched to this composite, which also results in an index of the global familiarity of the test item because the composite includes all the items.

The remainder of this section is organized as follows: Each model is described and followed by a brief tutorial showing how the central properties of the model are implemented. Each tutorial includes example calculations that illustrate how the model is applied to item and associative recognition. As will be seen, the relationship between item and associative recognition is central in these models and illustrates the mechanisms that underlie context effects. In addition to these tutorials, the models are also summarized in Table 1, in terms of their representational assumptions, how global matching is construed and implemented, how associations are represented, and how the model produces the intact advantage in intact–rearranged discriminations.

We begin by describing the models as they were originally proposed, describing more recent modifications as

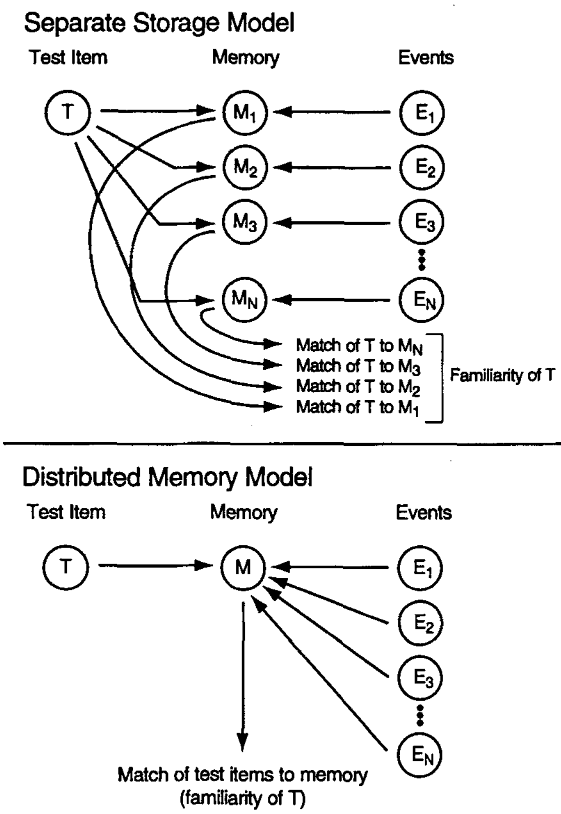

Figure 1. General schematic for global matching in separate (top panel) and distributed storage (bottom panel) memory models. To-be-learned events are labeled E_i, and their corresponding memories are M_i.

Table 1
Summary of Global Matching Models

Model	Representation	Global Familiarity Computation	Associations	Intact–Rearranged Discrimination
SAM	Separate cue-to-image strengths	Sum cue-to-image strengths	Interitem strengths $b > d$	Multiplicative cue combination
MINERVA 2	Vector, separate traces	Sum of dot products comparing test probe with each memory trace	Concatenation of item vectors	Nonlinear function relating similarity to activation
Matrix	Vector, distributed memory	Dot product comparing probe with distributed memory matrix	Multiplication of item vectors to produce matrix	Match of intact pair exceeds match of two partial pairs
TODAM	Vector, distributed memory	Dot product comparing probe with distributed memory vector	Convolution of item vectors	Match of convolution vector

they arise in the context of the data that motivate them. We do this because it is important to plot the evolution of these models as well as their current status. In addition, we follow the notation used in the original papers that introduced each model, rather than force the models into a common notational framework.

In our Application to Data section (below), we apply the models to the issues that were introduced earlier regarding effects of other items at study (list-length effects) and at test (context effects). These issues illustrate the two properties of global matching models that differentiate them from previous models—namely, that cues are combined interactively and that cues are matched to memory at a global level.

In the sections to follow, we will describe the basic models and apply them to relevant data. Some of the empirical results will pose difficulties for some of the models. It is important, before discussing the various problems, to discuss at a more general level how such problems should be interpreted. First, each of the models should be viewed as a framework within which specific models can be proposed. These frameworks evolve under the pressure of empirical data, and, although specific versions of models may be falsified by data, it is rare that the entire framework will be falsified.

Separate-Storage Models

The SAM Model

Each item in a study list is stored in a separate memory trace, called an *image*. An image is a unitized collection of information corresponding to some proportion of the information rehearsed and coded together in short-term store in conjunction with the word. According to SAM, memory is based on the strength of connection between the retrieval cues used to probe memory and these images in memory. To obtain predictions, one must specify the cues for a given recognition trial and specify the relationship between the cues and the images in memory. It is assumed that, for recognition, the cues consist of the test item(s) and experimental context. What remains is to specify the retrieval strengths between these cues and the images in memory.

The retrieval strength between a given cue (Q_i) and a given memory image (I_j) is denoted $S(Q_i, I_j)$, and is a function of their joint rehearsal in short-term store and preexperimental associations. Four parameters govern the strengths that are stored. The context cue has strength a to items on the list. The context is assumed to shift between lists such that the context cue for the current list has such a small strength of connection to previous lists that those images are effectively ignored. (For a detailed treatment of the role of context in recall, see Mensink & Raaijmakers, 1989.) An item cue has strength c to its own image in memory; this is called the *self-strength*. An item cue has strength b to images with which it was rehearsed, and it has strength d to images with which it was not rehearsed. The former is called *interitem* strength, the latter is *residual* strength and reflects the preexperimental strength of association between a cue and an image (i.e., the role of semantic memory). The only constraint on the values of the parameters is that $a, b, c > d$.

We can assume that a (b or c) units of strength are stored per second in the rehearsal buffer. Many factors, including the instructions to subjects, type of stimuli, simultaneous or sequential presentation, presentation time, and so on, influence the context strength stored as well as the relative amounts of interitem versus self-coding (i.e., the value of b vs. c). The value of d is also affected by a number of factors, including the semantic relationship between a studied target and an unstudied distractor (if DOG was studied, d for CANINE as a distractor would be greater than would d for UMBRELLA as a distractor).

Figure 2 shows the retrieval structure resulting from the study of a list that includes pairs AB and CD. Retrieval cues A, B, C, and D, a nonstudied item X, and the context cue are listed vertically, and images A, B, C, and D are listed horizontally. Each entry in the matrix represents the strength between a given cue and a given image in memory. Self-strengths (c) are shown on the diagonal; interitem strengths (b) are given for words studied together within a pair. Context strengths (a) are shown as the bottom row of the retrieval matrix, and all other cells contain residuals (d).

Recognition decisions are based on the total strength of the cues to all images in memory (functionally, the study list as demarked by connection to the context cue). Specifically, the familiarity for cues $Q_1, Q_2, \ldots, Q_M$ is given as

$$F(Q_1, Q_2, \ldots Q_M) = \sum_{j=1}^{N} \prod_{i=1}^{M} S(Q_i, I_j)^{W_i} \quad (1)$$

The equation shows that M cues are combined multiplicatively and summed over all N images in memory. To produce a response, the familiarity for a particular cue set is compared with a decision criterion. If the familiarity exceeds that criterion, an "old" response is given; otherwise, a "new" response is given.

Approximate[7] familiarities can be calculated using Equation 1 and the retrieval matrix in Figure 2. For example, the familiarity for a test of A, X, AB, and AD (probing with the test item[s] and context as cues) would be

$$F(A) = ac + ab + 2ad$$
$$F(X) = 4ad$$
$$F(AB) = 2abc + 2ad^2$$
$$F(AD) = 2abd + 2acd.$$

Familiarities for longer lists would simply add residual terms. For example, $F(A)$ for a 20-item list would be $ac + ab + (20 - 2)ad$. The variances of the familiarity may also be calculated as a sum of the variances for each cue-to-image strength (see Clark & Shiffrin, 1992; Gronlund & Elam, 1994). Once the means and variances are obtained, d' can be easily derived. Numerical predictions can be produced by selecting values for a, b, c, and d. For example, if we let $a = 1.0$, $b = c = 2.0$, and $d = 0.1$, $F(A) = 4.2$, $F(X) = 0.4$, $F(AB) = 8.02$, and $F(AD) = 0.8$. Several properties of the model are illustrated by this short example. The global nature of the familiarity computation is shown by the cue-to-image strengths being summed over all images in memory. Also, the interactive cue assumption is instantiated in the multiplicative combination of cues. It is this multiplicative cue combination

			Images			
			I_1	I_2	I_3	I_4
			A	B	C	D
	Q_1	A	c	b	d	d
	Q_2	B	b	c	d	d
Cues	Q_3	C	d	d	c	b
	Q_4	D	d	d	b	c
	Q_5	X	d	d	d	d
	Q_6	Context	a	a	a	a

Figure 2. Retrieval matrix for the SAM model following study of two word pairs, *AB* and *CD*. Item cues *A*, *B*, *C*, *D*, an unstudied item *X*, and context are listed on the abscissa, and the images for *A*, *B*, *C*, and *D* are listed on the ordinate. Entries in the matrix are the cue-to-image strengths.

that allows SAM to distinguish between intact (AB) and rearranged (AD) test pairs: $F(AB) > F(AD)$, provided that b and $c > d$.

MINERVA 2

Each item in the list is represented as a vector, where each cell in a vector is a random variable that may take on values of 1, −1, or 0, with equal probability. The features in the vectors might represent various features or attributes of the stimulus event. Let $\mathbf{w}_i(j)$ denote the jth cell of item i. Let $\mathbf{m}_i$ denote the representation of item i in memory. The information in cell j of $\mathbf{w}_i$ is stored in memory [i.e., $\mathbf{m}_i(j) = \mathbf{w}_i(j)$] with probability L; a value of zero is stored in element $\mathbf{m}_i(j)$ with probability $1 - L$. Performance increases with N, the number of cells in each vector (i.e., the dimensionality of the vector).

A word pair may be created by concatenating two vectors, each representing one of the component words. The familiarity of a test item (or test pair) is given by taking the dot product of the test-item or probe vector with each vector in memory [$\mathbf{m}_i(j)$]. The dot product or inner product gives a quantitative index of the match or similarity between any two vectors. The dot product is given by multiplying corresponding elements in each vector and summing the products. In MINERVA 2, the similarity (S_i) of the test probe $\mathbf{p}$ and a given trace in memory $\mathbf{m}_i$ is divided by N_i, where N_i is the number of features that are relevant to the comparison between $\mathbf{p}$ and $\mathbf{m}_i$ (the number of cells for which $\mathbf{p}(j)$ and $\mathbf{m}_i(j) \neq 0$). Thus,

$$S_i = \sum_{j=1}^{N} \mathbf{p}(j)\, \mathbf{m}_i(j) / N_i. \quad (2)$$

The activation of a given memory trace is a positively accelerated function of the similarity of that trace to the test probe. Specifically,

$$A_i = S_i^3. \quad (3)$$

The familiarity of the test probe is given by summing the activations for all items in memory,

$$F(\mathbf{p}) = \sum_{i=1}^{M} A_i, \quad (4)$$

where M is the number of traces in memory.

A numerical example, presented in Figure 3, illustrates the mechanics of the model. For simplicity, we set the encoding parameter L to 1.0. The top of the figure shows the vector representation for words A, B, C, and D, which are studied as pairs AB and CD, and X, which is a new item. The vectors representing each pair are simply the concatenation of the two vectors, and these are stored in memory. The bottom of Figure 3 gives the calculation for item (A and X) and associative recognition (AB and AD).

MINERVA 2 is like SAM in that both models calculate global familiarity by summing over all the traces in memory. MINERVA 2 sums similarity-based activations, whereas SAM sums cue-to-image strengths. In MIN-

<u>Memory</u>

<u>Items</u>

A	1	-1	-1	1
B	-1	-1	1	1
C	-1	1	-1	1
D	1	1	1	-1
X	1	1	-1	-1

<u>Studied as pairs, represented as concatenated vectors</u>

ab	1	-1	-1	1	-1	-1	1	1
cd	-1	1	-1	1	1	1	1	-1

<u>Item recognition</u>

$$F(A) = (1 + 1 + 1 + 1 + 0 + 0 + 0 + 0)/8 + (-1 + -1 + 1 + 1 + 0 + 0 + 0 + 0)/8$$
$$= (4/8)^3 + 0^3 = 0.125$$

$$F(X) = (1 + -1 + 1 + -1 + 0 + 0 + 0 + 0)/8 + (-1 + 1 + 1 + -1 + 0 + 0 + 0 + 0)/8$$
$$= 0^3 + 0^3 = 0$$

<u>Associative recognition</u>

$$F(AB) = (1 + 1 + 1 + 1 + 1 + 1 + 1 + 1)/8 + (-1 + -1 + 1 + 1 + -1 + -1 + 1 + -1)/8$$
$$= (8/8)^3 + (-2/8)^3 = .984$$

$$F(AD) = (1 + 1 + 1 + 1 + -1 + -1 + 1 + -1)/8 + (-1 + -1 + 1 + 1 + 1 + 1 + 1 + 1)/8$$
$$= (2/8)^3 + (4/8)^3 = .141$$

Figure 3. Numerical illustration for MINERVA 2.

ERVA 2, the interactive-cue assumption is instantiated by taking the dot product of the test probe and a memory vector. Intact–rearranged discrimination is due to the nonlinear function relating similarity to activation (see Equation 3). (This contrasts with SAM, where the intact advantage arises from the multiplicative combination of cues.) Test pair *AB* matches one memory trace exactly, whereas *AD* gives a half-match to two stored traces (*AB* and *CD*). Because of the nonlinearity, two half-matches are less than one exact match. [Note that without the cubing function, MINERVA 2 predicts $F(AB) = F(AD)$.]

Distributed-Storage Models

The Matrix Model

Like MINERVA 2, items are represented as vectors of elements in the Matrix model. Thus, items *A*, *B*, and *C* are represented by vectors **a**, **b**, and **c**. Associations between items are produced by multiplying their corresponding *N*-element vectors to produce an $N \times N$ outer product matrix. This matrix, denoted $\mathbf{ab}'$, represents the association between items *A* and *B*. The prime denotes a transposed row vector, to conform to the rules of vector multiplication. Hereafter, we will omit the prime, and the column and row vectors can be inferred from their order. The matrix representation has the important property that if the **ab** association matrix is premultiplied by **a**, the result is **b**, and if **ab** is postmultiplied by **b**, the result is **a**. Thus, the matrix representation satisfies what must be considered an essential property: either component can be used to retrieve the other. (Note that SAM and MINERVA 2 can accomplish this also.)

A list of word pairs is represented by summing together each of the association matrices. Thus, pairs *AB*, *CD*, and *EF* are represented in a single distributed memory **M**, where

$$\mathbf{M} = \mathbf{ab} + \mathbf{cd} + \mathbf{ef}. \tag{5}$$

Memory for a list of pairs is given by a single $N \times N$ matrix **M**, which is the sum of the matrices for each of the **k** studied pairs. Each cell in **M** is therefore equal to the sum of the corresponding cells in each of the study pair matrices.

Recognition decisions are made by taking the dot product of the test probe **p** and the single distributed memory matrix **M**. The probe **p** would correspond to the matrix **ab** for an intact pair and **ad** for a rearranged pair. The dot product is given by multiplying each cell in **p** by each corresponding cell in the memory matrix **M** and then adding these products. Thus, the familiarity of a test probe $\mathbf{p}$ is given as

$$F(\mathbf{p}) = \mathbf{p} \cdot \mathbf{M} = \sum_{j=1}^{N} \sum_{k=1}^{N} \mathbf{p}(j, k)\, \mathbf{M}(j, k), \tag{6}$$

where the indices j and k identify cell j, k in the matrix.

Figure 4 provides a numerical illustration of the model's properties. At the top of the figure are six vectors corresponding to the words A, B, C, D, E, and F. The vectors were constructed so that the expectation of each cell would be approximately $1/\sqrt{N}$. Because of this constraint, the sum of the squares of the vector elements is approximately 1.0 (see Humphreys, Bain, & Pike, 1989). In the middle of Figure 4 are the association matrices for each of the pairs, which are summed together to produce $\mathbf{M}$.

The test probe that is matched to memory must be of the same order as memory. Thus, if memory is an $N \times N$ matrix, then so must be each test probe. Matching test pairs to memory is straightforward; however, to probe memory with a single item, the N-element item vector must be converted into an $N \times N$ matrix. This is accomplished by multiplying the test-item vector with vector $\mathbf{r}$, whose elements are all fixed at $1/\sqrt{N}$. Thus, A as a probe would be matched to memory as $\mathbf{ar}$, and B as a probe would be matched to memory as $\mathbf{rb}$. The bottom of Figure 4 gives the dot product matches for item (A and X) and associative (AB and AD) recognition.

<u>Item Vectors</u>

a	.235	.408	.410	.781
b	.773	.523	.226	.279
c	.438	.692	.170	.548
d	.220	.270	.800	.489
e	.548	.170	.692	.548
f	.800	.489	.220	.270
x	.106	.840	.200	.490

<u>Association Matrices</u>:

ab

.182	.123	.053	.066
.315	.213	.092	.114
.317	.214	.093	.114
.604	.408	.177	.218

cd

.096	.118	.350	.214
.152	.187	.554	.338
.037	.046	.136	.083
.121	.148	.438	.268

ef

.438	.268	.121	.148
.136	.083	.037	.046
.554	.338	.152	.187
.438	.268	.121	.148

<u>Memory Matrix</u>: **M = ab + cd + ef**

.716	.509	.524	.428
.604	.483	.683	.498
.908	.599	.381	.384
1.163	.824	.735	.634

<u>Item recognition</u>
F(A) = **a r** • **M** = 2.495
F(X) = **x r** • **M** = 2.118

<u>Associative recognition</u>
F(AB) = **ab** • **M** = 2.464
F(AD) = **ad** • **M** = 2.072

Figure 4. Numerical illustration for the Matrix model. Prime symbols have been removed for clarity.

The numerical example illustrates the basic properties of the model: Global matching is due to summing events into a single composite memory at the time of study, rather than summing matches to separately stored traces at the time of retrieval. The interactive-cue assumption results from the dot product comparison between the test probe and the composite memory. Note that, in the Matrix model, for the study of word pairs, the events are associations, rather than items. The intact advantage is predicted because the match of an intact pair exceeds the match of two partial pairs.

TODAM

TODAM also represents items as vectors of elements. Each element is randomly sampled from a normal distribution with mean 0 and variance $1/N$. When a pair of words is studied, TODAM stores a vector corresponding to each studied word, plus another vector that represents the association between the words in the pair. Associations between words are formed, not by multiplication of item vectors, but by the *convolution* of the item vectors. The convolution of two vectors $\mathbf{a}$ and $\mathbf{b}$ (denoted $\mathbf{a} * \mathbf{b}$) is given as

$$(\mathbf{a} * \mathbf{b})_i = \sum \mathbf{a}_j\, \mathbf{b}_{i-j}, \tag{7}$$

where the index i denotes element i. Additional details about convolution can be found in Murdock (1982), and a graphic illustration can be found in Eich (1982).

TODAM assumes a distributed memory representation; therefore, item and convolution vectors are summed together into a single memory vector, $\mathbf{M}$. For a given pair, AB, item information is stored for A and for B, along with the convolution of A and B. Thus, memory for the pair is given by the sum

$$\mathbf{M}_1 = \lambda_1 \mathbf{a} + \lambda_2 \mathbf{b} + \lambda_3(\mathbf{a} * \mathbf{b}), \tag{8}$$

where the λ terms refer to weights on each vector. These weights reflect selective attention to components within the pair; they sum to 1.0 to reflect limited capacity in attention. This means that memory for the pair AB is a weighted sum of the item information for A, the item information for B, and the AB association.

As pairs are added to memory, information in memory may be forgotten. Thus, memory after presentation of pair j is given as

$$\mathbf{M}_j = \alpha\,\mathbf{M}_{j-1} + \lambda_1\,\mathbf{a}_j + \lambda_2\,\mathbf{b}_j + \lambda_3(\mathbf{a}_j * \mathbf{b}_j) \qquad (9)$$

where j is the serial position of the study pair, and α is a forgetting parameter. For values of α less than 1, memory is degraded as each pair is added to memory.

The familiarity values for A, X, AB, and AD are given by the dot product of the test probe with $\mathbf{M}$. Specifically,

$$F(A) = \mathbf{a} \cdot \mathbf{M}_j = \alpha^{j-1}\lambda_1$$
$$F(X) = \mathbf{x} \cdot \mathbf{M}_j = 0$$
$$F(AB) = \mathbf{a} * \mathbf{b} \cdot \mathbf{M}_j = \alpha^{j-1}\lambda_3$$
$$F(AD) = \mathbf{a} * \mathbf{d} \cdot \mathbf{M}_j = 0.$$

If probabilistic encoding is assumed, each of the above equations are simply multiplied by p, where p reflects the proportion of elements of the vector that are encoded at any presentation of the item. It is akin to MINERVA 2's L.

Figure 5 shows a numerical example for a list of three pairs. For simplicity, we set α, λ_1, λ_2, and λ_3 to 1.0. The item vectors are given at the top of the figure, followed by the three convolution vectors. The familiarities for item (A vs. X) and associative recognition (AB vs. AD) are given at the bottom of the figure. (The corresponding variances are tricky; Weber, 1988, and Murdock, 1992, give the component variances for a large number of probe sets.) Note that all vectors are centered at zero; hence, five-element item vectors can be added to nine-element convolution vectors by adding the required number of flanking 0s to the item vectors.

TODAM, like the Matrix model, produces a global match by matching a test probe to a composite memory, which includes all items from the list (hence, global matching). TODAM also instantiates interactive cuing; A and B are combined by convolution and then matched against memory. The intact advantage is predicted due to the match of the convolution vector.

Application to Data

In this section, we present recognition data that bear on the two fundamental properties that define this class of models. Experiments on context effects and experiments that examine the relationship between item and associative recognition are discussed in terms of interactive-cuing assumptions. Global matching assumptions are discussed with respect to list-length, list-strength, item-strength, and global similarity effects.

Interactive-Cue Combination

A defining characteristic of global matching models is that multiple cues are combined interactively. This property allows the models to produce context effects and distinguish intact from rearranged test pairs, and it defines how the models represent associative information. In this section, we will further explore how cues are combined in each model, how each model represents associative information, and how the assumptions of each model stand up to data.

Context Effects

To begin, we discuss how the models account for the facilitatory effects of reinstating verbal context at the time of test. In a typical experiment, words are studied as pairs, *AB*, *CD*, *EF*, and a word (e.g., *B*) is tested alone (**_B**), in a different context (*C***B**), or with the context item it was presented with at study (*A***B**). The boldface indicates that the subject's task in all cases is to make a recognition decision to **B**. The test item **B** is better recognized when it is presented at test with the original verbal context (Clark & Shiffrin, 1992; Humphreys, 1976; Light & Carter-Sobell, 1970; Thomson, 1972; Tulving & Thomson, 1971).

With the exception of the SAM model, the models produce context-based facilitation because associative information contributes to the match of the test probe only when *B* is tested with *A*, but not when *B* is tested alone or with different context items. If cuing was not interactive, there would be no difference between *A***B** and *C***B** (both *A* and *C* were previously studied). It is this same property that allows the models to distinguish between intact and rearranged test pairs.

It is important to distinguish the associative account of context effects from an account given within the framework of local match models, which goes by such names as encoding bias, meaning selection, or selective activation (e.g., Light & Carter-Sobell, 1970). The encoding-bias explanation of context effects is nonassociative and, therefore, fundamentally different from the account offered within the framework of global matching models. According to this account, context functions to bias the encoding of the test item, and, thus, when the context is changed at test, the test item is encoded differently. Because the encoding at test does not match the encoding at study, the familiarity of the test item will be decreased. Bain and Humphreys (1988; Humphreys, 1976; Humphreys & Bain, 1983) and Clark and Shiffrin (1992) have argued against the encoding-bias hypothesis. The arguments are somewhat long, and, rather than reproduce them here, we refer the reader to the original papers.

Associative account of context effects. In the Matrix model, for study of word pairs, information is stored at the level of pairs; thus, when *A***B** is tested, a pair is matched, while **_B** produces only a (small) partial match to a pair. TODAM produces context facilitation in a manner similar to the Matrix model. In their application of the model, Clark and Shiffrin (1992) assumed that for recognition of *A***B**, the information in the probe was given by $\mathbf{b} + (\mathbf{a} * \mathbf{b})$. For recognition of **_B**, the probe consists of only $\mathbf{b}$. The match for the *A***B** probe will be greater than the match of the **_B** probe because of the match of the associative information $\mathbf{a} * \mathbf{b}$. MINERVA 2 is like the Matrix model in that similarity is computed over word pairs, rather than items. The activation resulting from two words in a pair exceeds the activation of two words not in a pair because of the nonlinear similarity function: the match for *A***B** is more than twice the match of **_B** alone.

Item Vectors

a	-0.048	0.250	-0.464	-0.186	-0.637
b	0.213	0.299	0.247	0.154	0.934
c	0.568	-0.338	-0.511	-0.129	-0.282
d	0.351	-0.155	0.170	-0.581	0.074
e	-1.081	0.564	-0.442	-0.374	0.128
f	-0.528	0.564	-0.442	-0.381	-0.073
x	1.044	0.305	0.317	-0.157	-0.607

Associations (by convolution)

a*b	-0.010	0.039	-0.036	-0.124	-0.313	-0.074	-0.619	-0.272	-0.595
c*d	0.199	-0.207	-0.030	-0.353	0.073	0.293	-0.011	0.154	-0.021
e*f	0.571	-0.369	0.367	0.529	-0.188	0.169	0.163	-0.021	-0.009

M = **a** + **b** + (**a*b**) + **c** + **d** + (**c*d**) + **e** + **f** + (**e*f**)

0.760 -0.537 -0.224 0.738 -1.517 -1.107 -0.323 -0.139 -0.625

Item recognition

F(A) = a•M = 1.310

F(X) = x•M= -0.119

Associative recognition

F(AB) = (a*b)•M = 1.054

F(AD) = (a*d)•M = 0.280

Figure 5. Numerical illustration for TODAM. Let $\lambda_1 = 1, \lambda_2 = 1, \lambda_3 = 1, \alpha = 1$.

TODAM, MINERVA 2, and the Matrix model pro-uce context effects because of the associative informa-on in the compound cue (*A***B**) that is not present in the em cue (_**B**) or in the rearranged cue *C***B**. However, in AM, associative information is accessed by the single em cue. Consequently, SAM predicts a somewhat dif-erent pattern of results than do other models. For SAM, is useful to distinguish between removed context (_**B**) nd changed context (*C***B**). SAM predicts that *A***B** is su-erior to *C***B**. However, SAM does not predict an advan-ge for *A***B** relative to _**B**.

The reason lies in the attention weights on the cues, hich are constrained to sum to 1.[8] Any weight given to e context word *C* must be taken away from the weight iven to the test item **B**. The problem can be described as cost–benefit tradeoff, where the costs are attentional nd the benefits are associative (Clark & Shiffrin, 1987, 992). For context-based facilitation to occur, the costs f removing weight from the test item must be more than ffset by giving that weight to the context word. It must dd information to the memory probe that is not given by e test item. Presumably, that added information would e associative; however, in SAM, the association linking *B* to *A* is already part of the sum that results from the familiarity for _**B**. Thus, the new information that the context word *A* could add to a test of *B* is already entered into the familiarity sum for _**B**. Thus, shifting weight from the test item *B* to the context *A* does not bring additional information.

Changed context effects (*A***B** better than *C***B**) are quite reliable; however, results concerning removed context (*A***B** vs. _**B**) are more complicated. Some studies have shown a removed-context deficit (_**B** worse than *A***B**; Clark & Shiffrin, 1992; Humphreys, 1976, 1978; Tulving & Thomson, 1971); others have shown either no difference or even a small advantage for the removed-context condition (Clark & Shiffrin, 1987, 1992; DaPolito, Barker, & Wilant, 1972; Gillund & Shiffrin, 1984; Gronlund, 1986).

Results from an experiment by Clark and Shiffrin (1992) suggest that the variability in the results of the removed-context condition is due to the formation of associations. They showed a removed-context advantage for low-frequency words but a same-context advantage for high-frequency words. Because it is difficult to form associations using low-frequency words, the attention given to the cue is wasted because the cue cannot combine with the test item to retrieve associative information.

The strengths and weaknesses of SAM relative to other models are complementary. SAM cannot predict an advantage for *A***B** over _**B**, but the other models cannot predict an advantage for _**B** over *A***B**.

Variation in Environmental Context

In this section, we discuss another type of context shift. Here, context is defined as external background context, or environmental context. In the standard paradigm, subjects learn a list of words in one room and are tested in either the same room or a different room. With a few notable exceptions (see Bjork & Richardson-Klavehn, 1989; Fernandez & Glenberg, 1985), this kind of context shift produces a decrement in recall performance but typically does not produce overall differences in performance for recognition (Smith, 1988; Smith, Glenberg, & Bjork, 1978).

This is not to say that shifts in background context have *no* effect on recognition. Murnane and Phelps (1993, 1994) manipulated background context in terms of varying colors and locations of study words on a computer screen. In their experiments, a shift in background context led to a decrease in both hit and false-alarm rates. Furthermore, they showed that all global matching models can produce this pattern of results because, in general, context shifts decrease the familiarity of both targets and distractors. Moreover, not only must the models predict decreases in both hit and false-alarm rates due to context shifts, they must predict that the decreases are parallel, thus producing no difference in overall performance. Can the models produce this pattern of results?

As a prelude to addressing this question, we should make a clearer connection between environmental and verbal context experiments. In the standard environmental context-shift experiment, words *A*, *B*, and *C* are studied in a context, *S*, and targets (*A*) and distractors (*X*) may be tested in the same context or a new context, *N*. The test word–context pairs may be denoted *A–S* (target–same context), *A–N* (target–new context), *X–S* (distractor–same context), and *X–N* (distractor–new context). Again, the typical result is that target–distractor discrimination does not differ for same- and new-context testing.

How do these results compare with the corresponding verbal context shift? The test word–context pairs here would be denoted *A–B*, *A–Y*, *X–B*, and *X–Y*, where *A* and *B* were studied together, and *X* and *Y* are new words. The relevant test items (although not previously identified as context and test item) have been tested in several experiments, the results of which show a sizable context-shift effect for targets but a smaller (Humphreys, 1976; Tulving & Thomson, 1971) or zero (Clark, 1992) context-shift effect for distractors.

These results suggest that, compared with verbal context, environmental context is either stored differently or utilized differently as a cue (or both). Otherwise, context shifting would show the same results, irrespective of the kind of context that is shifted. On the basis of their results, Murnane and Phelps (1993, 1994) proposed that item information and context are combined in additive fashion. However, previous treatments of environmental context in global matching models have treated environmental context like verbal context and combined item information and context in a multiplicative fashion. Humphreys, Bain, & Pike (1989) suggested the additive-cue possibility within the Matrix model but did not instantiate the assumption in the model.

TODAM, CHARM, and the Matrix model. Metcalfe and Murdock (1981) simulated free recall in a model that would later evolve into CHARM and TODAM. This model represented environmental context as a vector, which was convolved with the first few items on the list (subsequent list items were convolved with each other). This would incorrectly produce a d' difference for recognition of the first few list items and no effect of context for subsequent list items.

Similarly, within the Matrix model, Humphreys, Bain, and Pike (1989) also incorporated environmental context as a vector that is associated to list items in the same manner as associations are formed for words—in this case, by outer product multiplication. The problem is the same as that for the Metcalfe–Murdock model: shifts of environmental context will produce decrements that parallel shifts of verbal context, contrary to the data. Thus, environmental context must not be associated to items by convolution (TODAM, CHARM), or by outer product multiplication of a context vector with a word vector (Matrix model).

One can produce the Murnane and Phelps results by assuming that environmental context and item information combine additively but do not form an interactive cue. For example, in TODAM, a memory probe that is the sum of context vector **c** and item vectors **a** or **x** will produce parallel effects on targets and distractors. However, if the item–context convolution is included, target effects will be larger than distractor effects.

MINERVA 2. In MINERVA 2, environmental context could be represented as a vector that is concatenated with the word vectors. However, the cubing function introduces a nonlinearity that will produce an overall performance decrement for shifted-context testing. If the cubing function is removed, MINERVA 2 can correctly predict that a context shift will have parallel effects on targets and distractors and thus no overall effect on d'. However, removing the cubing function would eliminate MINERVA 2's ability to discriminate intact from rearranged test pairs. However, it is possible to propose that the cubic function might apply only to word associations, not to context. Thus, for pair *AB* in environmental context *C*, the activation of the *AB–C* trace would be $S_{ab}^3 + S_c$, where subscripts denote pair (*AB*) and context (*C*) portions of the trace vector.

SAM. The SAM model correctly predicts that environmental context shifts will not produce changes in the level of performance. In fact, SAM makes identical predictions for shifts in verbal or environmental context, the prediction being that such shifts will have no effect on the overall level of performance. Thus, the error in prediction for SAM is the inverse of the error made by the other global matching models: SAM correctly predicts no effect of environmental context on d', but incorrectly predicts no effect of verbal context.

To summarize, all of the models except SAM can incorporate environmental context in a manner that is different from how they incorporate verbal context and, in so doing, can produce a different pattern of results for environmental and verbal context shifts. A problem that remains is how to treat environmental context in recall. In some cases, the fix for recognition may be problematic for recall. For example, in TODAM, if the context–item convolution ($\mathbf{c} * \mathbf{a}$) is not stored, context will be a useless cue in recall. Since discussions of recall are beyond the scope of this paper, we will not pursue this problem further.

Associative and Item-Specific Information

Global memory models account for item and associative recognition with a single retrieval mechanism (although different processes may be required to form the cues). This contrasts with associative network models (Anderson, 1972; Anderson & Bower, 1973), in which associative information serves only as a *pathway* for directing search from one node to another. In such models, associative information is not retrieved directly but rather provides a means to retrieve item-specific information stored at a given node. Thus, associative information would have no role in recognition.

The assumptions of global matching models vary regarding the relationship between associative and item information. We address two questions: First, are item and associative information stored as *part of the same representation*? As a consequence, strengthening item information necessarily implies strengthening of associative information (and vice versa). Second, do item and associative information *make independent contributions* at retrieval? Independent contributions are demonstrated by experimental manipulations that affect one kind of recognition, but not the other. These representation and retrieval issues are asymmetrically related in global matching models. If item information and associative information are part of the same representation, they will show a dependency at retrieval as well. However, separation at storage does not imply independence at retrieval. We elaborate below, for each model.

MINERVA 2. In MINERVA 2, word pairs are stored as one single vector, which is the concatenation of the two component vectors. Associative information results from the fact that A and B are stored as a single memory trace, and item-specific information is a subset of that memory trace. Thus, it is clear that item-specific and associative information are part of the same representation and, consequently, make nonindependent contributions at retrieval. Manipulations that would increase the strength of item information will increase performance for associative recognition, and vice versa.

The Matrix model. In the Matrix model, both item information and associative information reside within each association (like MINERVA 2), so there is no distinction between the two kinds of information at the level of representation. Furthermore, because item information is derived from the stored associative information, the model does not allow distinct contributions of item and associative information at retrieval. In other words, the item match and the pair match are based on overlapping sets of elements in memory. Specifically, the item match is based on a subset of the pair-match elements.

SAM. SAM explicitly distinguishes between item and associative information at the level of representation. Associative information corresponds to the interitem strength parameter b; item-specific information corresponds most closely to the self-strength parameter c (although one might also consider contextual information to be item-specific in the sense that it represents a connection between context and a particular item). However, item information and associative information do not make distinct contributions to retrieval. The familiarity computation for A (where A was studied in pair AB) will sum over the A-to-B association; thus, associative information contributes to item recognition. Similarly, self-strength (item-specific information) contributes to the familiarity difference between intact and rearranged pairs in associative recognition.

TODAM. TODAM distinguishes between item-specific vectors and associative (convolution) vectors and, consequently, distinguishes between item and associative information at the level of representation (despite summing both together into a single distributed memory vector). TODAM allows distinct contributions from item and associative information at retrieval: Item recognition is based on item-specific information, and associative information is largely irrelevant (it contributes only noise); similarly, associative recognition depends on associative information, and item-specific information is largely irrelevant.

In sum, MINERVA 2 and the Matrix model represent item and associative information in the same representation, and the two kinds of information do not make independent contributions to recognition. SAM represents item and associative information separately but retrieves them together. TODAM distinguishes between item and associative information at the level of representation and at retrieval. In the next section, we examine data relevant to the relationship between item and associative information.

Relevant Data

Until recently, item information and associative information have been studied using recall procedures (Hovland & Kurtz, 1952; Hunt & Einstein, 1981; Postman, 1962; Underwood, Runquist, & Schulz, 1959); in other cases, item information has been examined using recognition, and associative information has been examined using cued or free recall.

More recently, associative information has been examined using the associative recognition task, which requires discrimination of intact from rearranged test pairs. Like cued recall or paired-associate learning, associative recognition requires subjects to know which words appeared *together*. However, like the more standard item recognition procedures, associative recognition does not require the recovery and output of stored information. Different patterns of results are shown for item and associative recognition, across a variety of experimental manipulations at study and test. Some of these results are summarized below.

Effects of encoding instructions. Item recognition and associative recognition show very different patterns of results, depending upon instructions given at the time of study. Whereas previous results have shown that frequency judgments for items are quite accurate under incidental instructions (Hasher & Zacks, 1979), frequency judgments for correct combinations of items are quite poor under incidental instructions (Greene, 1990). Hockley and Christi (in press) have shown that associative recognition performance decreased when subjects expected an item recognition test, but item recognition performance did not change when subjects expected an associative recognition task. These results suggest that item-specific information may be stored relatively well with a variety of study strategies, but that the formation of associations requires explicit instructions to form associations.

One cannot be sure of the rehearsal strategies subjects might be using to produce this pattern of results, and it is therefore useful to examine studies in which subjects were given explicit instructions regarding encoding of pairs. McGee (1980) and Bain and Humphreys (1988) instructed subjects to study pairs of words using either interactive or separate images. The results showed a crossover interaction: item recognition was better using separate imagery, whereas associative recognition was better using interactive imagery. These data are a challenge to models for which item information and associative information are part of the same representation.

Finally, there is evidence (Yuille & Paivio, 1967) that subjects will use mediators to interrelate a pair of words (HOUSE–ROPE; "The old HOUSE had *doors* hinged with a ROPE"). Alternatively, subjects may try to learn a pair by relating them using imagery (e.g., Bower, 1972; e.g., SCISSORS–WHEAT; "large shears harvesting wheat"). The models assume that associations do not differ qualitatively at retrieval, despite what is done at storage. As a result, the simplifying assumption is made that these qualitatively different associations can be expressed as quantitative differences (i.e., higher or lower strengths, more or less complete encoding).

Time course of retrieval. The most straightforward extension of a single global matching process into the temporal domain would result in the simultaneous availability of item and associative information. Gronlund and Ratcliff (1989) examined this prediction using the signal-to-respond procedure (Reed, 1973). On each test trial, the subject is presented with a test item followed by a signal to respond "old" or "new." The subject must make a response immediately following (but not prior to) the presentation of the signal. An important feature of the signal-to-respond procedure is that it can be used to identify the point at which retrieved information first becomes available. At very short lags, before information is available, subjects can only guess, and performance is at chance. Once information becomes available, the performance function begins to rise above chance. Gronlund and Ratcliff (1989) showed that the performance function for item recognition rises above chance approximately 150 msec earlier than the performance function for associative recognition. This presupposes that these two kinds of information make distinct contributions to recognition, which is problematic for models that treat item and associative information inseparably at retrieval. Moreover, the false-alarm rate functions for rearranged pairs (*AD*) are nonmonotonic; as signal lag increases from the shortest lags, false-alarm rates initially increase but then decrease as the signal lag is increased further.

Word-frequency effects. Item recognition and associative recognition are further dissociated with respect to the effects of word-frequency manipulations. Previous research shows that recall performance is better for common, high-frequency words than for rare, low-frequency words, but item recognition performance is better for low-frequency words than for high-frequency words (see Gregg, 1976, for a review). Recent studies have shown a similar dissociation within recognition; associative recognition shows a high-frequency advantage (Clark, 1992; Clark & Shiffrin, 1992) or shows no effect of word frequency (Hockley, 1994), contrary to the low-frequency advantage shown for item recognition.

Similarity effects. Similarity effects also are quite different for item and associative recognition. In a forced-choice item recognition task, Tulving (1981) and Hintzman (1988) showed better performance when the distractor on a given trial was similar to the target than when the distractor was similar to some other item in memory (but not the target on that trial). Given study of items A, B, and C, forced-choice recognition was better for A versus A' (where A' is a new item similar to A) than for A versus B' (where B' is a new item similar to B, but not similar to A). Although this similar-distractor advantage is somewhat counterintuitive, it is predicted by all of the global matching models which assume vector representation of items (CHARM, TODAM, Matrix, and MINERVA 2). These models produce the similar distractor advantage because the familiarities of similar test items are correlated, and thus the variance of the difference distribution for target and distractor is smaller (see Clark, Hori, & Callan, 1993, and Hintzman, 1988, for additional details).

Clark et al. (1993) have shown that all of the global matching models predict a similar-distractor advantage for associative recognition, also due to correlated familiarities. (Note that TODAM can also predict no difference, depending on parameters.) To test the correlated-familiarities hypothesis for associative recognition, Clark et al. presented subjects with study pairs, *AB*, *CD*, *EF*, *GH*, *IJ*, and so on, and tested forced-choice recognition with similar or dissimilar distractors. They used the term *OLAP* to describe test trials with similar or overlapping test pairs and the term *NOLAP* for dissimilar, or nonoverlapping, test pairs. An OLAP test trial would test the target *AB* against distractors *AD* and *AF*, whereas the NOLAP test trial would test *AB* against *CF* and *GJ*. Contrary to all of the models, and contrary to the item recognition results of Hintzman, Clark et al. showed a reliable NOLAP advantage (a dissimilar-distractor advantage).

Differential forgetting. Hockley (1991, 1992a; Murdock & Hockley, 1989) examined the forgetting of item and associative information as a function of study–test lag in a continuous recognition procedure and a study-test procedure. Item recognition performance decreased as the number of items presented between study and test increased. However, for associative recognition, d' was constant as a function of the number of items intervening between study and test. It should be noted that, irrespective of how the item–associative relationship may be represented, these results are difficult for any model to explain because there appears to be no forgetting of associative information as a function of study–test lag (at least up to test lags of 24 pairs).

Summary and modifications to the models. Taken together, these results are best explained by assuming that item information and associative information are not part of the same representation and that they make distinct contributions to retrieval. The crossover interactions shown for imagery instructions (Bain & Humphreys, 1988; McGee, 1980) and for word frequency (Clark, 1992; Clark & Burchett, 1994) cannot be produced by models that do not distinguish between item and associative information at the level of representation (i.e., MINERVA 2 and the Matrix model). In addition, Clark (1992) derived d' expressions showing that SAM could not decouple item and associative recognition for word frequency, even though SAM does distinguish between item and associative information at the level of representation. The problem, of course, is that both kinds of information are retrieved together. Presumably, these problems would also hold for the McGee results and the Bain and Humphreys results. Thus, the only model that might account for these interactions is TODAM, because item information and associative information are decoupled both at the level of representation and at retrieval.

None of the models, including TODAM, can account for distractor-similarity effects in associative recognition (Clark et al., 1993). Thus, none of the current models in their present configurations fully account for the relationship between item and associative information. We next explore three modifications to the current models that might help address the mispredictions: (1) different cues are responsible for item and associative recognition, (2) associations are stored as higher order units, distinct from their component items, and (3) different processes are required for item and associative recognition, with associative information requiring recall. Note that this last option is contrary to recognition being based on a single process.

Different cues. SAM and MINERVA 2—two models that do not allow distinct contributions from item and associative information—could be modified so that two cues are required for recognition. In MINERVA 2, memory could be probed with the component-item vectors as well as the concatenated-pair vector. In SAM, memory could be probed with the component items as cues (what Gronlund & Ratcliff, 1989, called *concurrent* cues) and the two items as an interactive cue (as is normally done). This modification might allow the models to address the time-course differences, assuming that creating the concatenated or interactive cue would take additional time. However, even with this modification, the models still do not address the differential effect of word frequency, distractor-similarity effects, or differential forgetting.

Higher order associations. A second means whereby item and associative information might be dissociated is by assuming that associations are stored as higher order units that can contribute to recognition independently of the items that are part of the association. In their original form, CHARM and TODAM exhibit this property. In CHARM, items are stored as autoconvolutions (i.e., $\mathbf{a} * \mathbf{a}$ and $\mathbf{b} * \mathbf{b}$), and associations are stored as convolution vectors. In TODAM, items are stored as vectors, and associations are stored as convolution vectors. The association formed by $\mathbf{a} * \mathbf{b}$ is independent of both $\mathbf{a}$ and $\mathbf{b}$. Put another way, when $\mathbf{a}$ and $\mathbf{b}$ are associated, the result is a memorial unit that is unlike either $\mathbf{a}$ or $\mathbf{b}$. As a result, item and associative recognition performance can vary independently.[9]

The Matrix model can be modified to behave in a similar fashion. If items are stored as autoassociations (i.e., $\mathbf{aa}'$ and $\mathbf{bb}'$), then item and associative information will be independent in the Matrix model also. It is interesting to note that, in the Matrix model, item–associative independence is not achieved by changing assumptions about associations but rather by changing the assumption about storage of item information.

The common denominator in CHARM, TODAM, and the Matrix model is that associations are represented as units rather than as links between units. The formation of associative units might require time and effort, which could account for the 150-msec difference in initial availability in the response-signal results (Gronlund & Ratcliff, 1989) and the incidental–intentional memory differences (Greene, 1990).

Higher order associations have been incorporated into recent applications of SAM (Clark, 1995; Gronlund, 1986; Mensink & Raaijmakers, 1989; Shiffrin, Murnane, Gronlund, & Roth, 1988). Higher order units can be incorporated in SAM for both cues and images. That is, *AB* can be stored in memory in addition to the item-level images *A* and *B*, and memory can be probed with the higher order cue *AB* in addition to the item-level cues *A* and *B*. One can create such a model by simply adding the *AB* unit as both a cue and a memory trace. This would add another level of cue-to-item retrieval strengths: item-level cues to higher order associations, higher order cues to item-level traces, and higher order cues to higher order traces. Item–associative independence could be achieved by assuming that item-to-association and association-to-association retrieval strengths are independent.

One potential problem facing the higher order version of SAM is that it has *two* ways of storing associative information: as *A*-to-*B* connections and as *AB* units. It may be that both kinds of associations are stored—in which case, the task is to specify the conditions that underlie the dual representations of associations and how the two kinds of associations interact at storage and retrieval. On the other hand, it may be that item-level associations are simply superfluous. However, the redundancy problem may be more apparent than real. Because retrieval strengths are typically presented in a cues-by-traces matrix, it *appears* that the item-level associations are obligatory in the model and a mandatory component of the familiarity as it is summed across the row in the matrix. However, the matrix is only a convenient way of displaying cue-to-image strengths, not a structure of the conceptual model.

The higher order version of SAM effectively separates item and associative information in a manner quite similar to that of TODAM, CHARM, and the modified version of the Matrix model. In all of these models, memory would be probed with a weighted combination of item-level and association-level cues, which would be matched to item-level and associative-level components in memory. Such a model may have the capability to handle the item–associative data, although additional theoretical work would be necessary.

It is not clear how one might implement something like higher order units in MINERVA 2. In TODAM, CHARM, and the modified Matrix model, associations are stored in addition to the component items, which is akin to a higher order unit. In MINERVA 2, associations are not stored in memory but rather emerge from memory at the time of retrieval as a result of the cubing function. One could somehow produce associations and store them in addition to items; however, this would essentially transform MINERVA 2 into a separate-trace version of TODAM and would make the cubing function superfluous.

Contribution of recall. Another means of dissociating item and associative recognition would be to assume that associative recognition decisions are based in part on recall-like retrieval. The addition of a recall-like retrieval component has the potential to handle word-frequency and distractor-similarity effects shown in associative recognition.

Like recall, associative recognition often shows a high-frequency–word advantage (Clark, 1992; Clark & Burchett, 1994; Clark et al., 1993). Like recall, associative recognition shows large performance differences comparing incidental and intentional instructions (Greene, 1990; Hockley & Christi, in press), and associative recognition performance increases when interactive imagery instructions are given at study (Bain & Humphreys, 1988; McGee, 1980). Also, in the response-signal procedure (Gronlund & Ratcliff, 1989), the delay in the rise in the d' function for associative recognition may reflect the operation of a slower recall process.

Recall may also underlie the NOLAP advantage for associative recognition shown by Clark et al. (1993). Recognition decisions can be made by using individual words as cues to recall correct study pairings. For example, subjects might recognize *AB* as intact by using *A* to recall *B*, or they may reject *CF* by using *C* to recall *D*. Why should this kind of a cued-recall strategy produce better performance for NOLAP test pairs? The reason is because there are more cues in the NOLAP test condition (*A*, *B*, *C*, *F*, *G*, and *J*) than in the OLAP test condition (*A*, *B*, *D*, and *F*) and, thus, more independent probes of memory.

Nobel and Huber (1993) verified this logic within SAM. However, even with recall modifications, TODAM and MINERVA 2 showed no difference for OLAP and NOLAP test conditions. Of course, there are many ways in which cued recall could operate in associative recognition, and a different implementation of recall might well produce a NOLAP advantage in these models.

Nobel and Huber's simulations with SAM demonstrate that the NOLAP advantage can be produced by a cued-recall strategy. Clark and Hori (1995) have provided additional empirical support for the recall-based explanation of the NOLAP advantage. They reasoned that if the NOLAP advantage is produced by recall, then under conditions where recall is difficult (and thus its contribution to recognition decisions is minimal), the NOLAP advantage should be eliminated. Consistent with this, the NOLAP advantage disappeared when subjects studied very long lists consisting of 100 pairs but replicated the NOLAP advantage with a short list of 24 pairs.

There are dissociations between cued recall and associative recognition that may be taken as evidence against the operation of a recall-like retrieval mechanism. In an experiment by Dyne, Humphreys, Bain, and Pike (1990), subjects studied word pairs such as *AB* and *EF*. For half of these pairs, the words were presented later in the list in other pairs (i.e., *AD*, *CB*, *EH*, and *GF*); for the other half of these pairs, the words were not repeated. In both cases, memory was tested with intact (*AB*) and rearranged (*AF*) test pairs. Dyne et al. derived predictions that showed that the global matching models predict interference in cued recall but no interference in associative recognition (akin to the Hockley, 1991, 1992a, differential forgetting data). The results showed exactly this pattern, consistent with the global matching models *without* a recall component. Similarly, Clark and Burchett (1994) showed that the high-frequency advantage for associative recognition remains when mixed-frequency lists are studied, whereas the advantage disappears in cued recall.

It seems reasonable that recall processes may operate to a lesser or greater degree in associative recognition, depending upon many factors. For example, recall might be limited in the Dyne et al. (1990) study because the study lists were very long. The NOLAP advantage (thought to be due to recall) disappeared when subjects studied very long lists. It is also possible that not all components of recall may be necessary for associative recognition. A subject presented with the intact target *AB* may "recall" *B* using *A* as a cue but be able to bypass aspects of information recovery and output because *B* is already available in the cue.

Much of the data reviewed in this section are compatible with a two-process model in which global matching is augmented by recall-like retrieval processes. If one assumes that recall processes operate in associative recognition, it seems absurd to assume that they operate exclusively in associative recognition and never in item recognition (although how recall may operate in these two tasks may be quite different). The broader contribution of recall in recognition is an issue we will explore further in the next section.

Global Matching

The second defining property of global matching models is that the test probe is matched to memory at a global, rather than local, level. The probe is matched to a memory consisting of multiple events (which may have separate or composite representation) that are accessed in parallel. In evaluating the global match assumption, the relevant experiments are those in which characteristics of other items in memory are varied. To this end, we will discuss list-length, list-strength, and global-similarity effects.

List-Length Effect

The list-length effect shows that, as the number of items on a list increases, recognition performance decreases. The same result is observed in recall (Murdock, 1962). SAM, the Matrix model, and MINERVA 2 produce recognition list-length effects by a common mechanism: the variances of both target and distractor distributions increase as items are added to the list, because each item added to memory contributes an additional source of noise to the matching process. Increases in list length add to the variance, but the difference between the means of the familiarity distributions remains constant, thus producing a decrease in performance.

List-length effects can also be produced by other mechanisms (see Murdock & Kahana, 1993). For example, in TODAM, setting the forgetting parameter $\alpha < 1$ will produce a performance decrement for longer lists because earlier items will be forgotten as each item is added to the long list. The forgetting account of the list-length effect implies that if the study–test lag is equated for short and long lists, the list-length effect will disappear. However, experiments that control for study-test lag continue to show a list-length effect (Bowles & Glanzer, 1983; Gronlund & Elam, 1994; Murnane & Shiffrin, 1991a).

Gronlund and Elam (1994) tested the increasing variance account of the list-length effect using receiver operating characteristic (ROC) analyses. ROC functions can be used to estimate the ratio of the distractor-to-target variance of the underlying familiarity distributions. If each item on the list adds noise to the matching process (as assumed by SAM, MINERVA 2, and the Matrix model), the variance ratio of distractors to targets should approach 1.0 as the list length gets longer (because both variances become more nearly equal).

Logic of the ROC analysis. Subjects make recognition decisions based on a 6-point confidence scale. The scale ranges from *sure yes* (*I'm sure I've studied the test item*) to *sure no*, with *likely yes*, *guess yes*, *guess no*, and *likely no* responses in between. To make these decisions, it is assumed that the subject sets five response criteria, dividing the underlying familiarity scale into six regions (see Figure 6A). An ROC function results from plotting the proportions of hits versus the proportion of false alarms at each response criterion (see Figure 6B). If the underlying target and distractor familiarity distributions are normal (as predicted by the models), then a z transformation of the ROC function will produce a linear ROC (Figure 6C). The intercept of the transformed-ROC function provides an estimate of d'; the slope is an estimate of the ratio of the distractor-to-target standard deviations. This ratio will be less than 1.0 if the target variance is greater than the distractor variance, and it will be greater than 1.0 if the distractor variance is larger.

In an experiment where multiple lists were studied and tested, Gronlund and Elam (1994) found a list-length effect (the intercept was less in the long-list condition), but the slope (the standard deviation ratio) was constant, and approximately 1.0, as a function of list length. This is contrary to the predictions of SAM, MINERVA 2, and the Matrix model. The slope did increase with list length when a single list was studied and tested, although the models predicted a larger increase than was observed.

Evidence for recall. Yonelinas and Jacoby (1994) examined list-length effects within the framework of a two-process (recall-plus-familiarity) model and argued that list-length effects are produced exclusively by the recall component and that the contribution of familiarity remains constant as list length increases. Their analyses are based on an inclusion–exclusion process-dissociation procedure (Jacoby, 1991). The logic of the inclusion–exclusion comparison is as follows: Subjects study two sets of items, set *A* and set *B*. In the include condition, subjects are to respond positively to items in set *A* and set *B* and respond negatively only to new items. In the exclude condition, subjects respond positively only to set *A* items and negatively to set *B* items and to new items. The probabilities for positive responses are a combination of recall *R* and familiarity *F*. The probability of a positive response in the include condition is

$$P(\text{include}) = R + F - RF,$$

where *R* is the probability of recall, *F* is the probability that the familiarity is above criterion, and *RF* is their in-

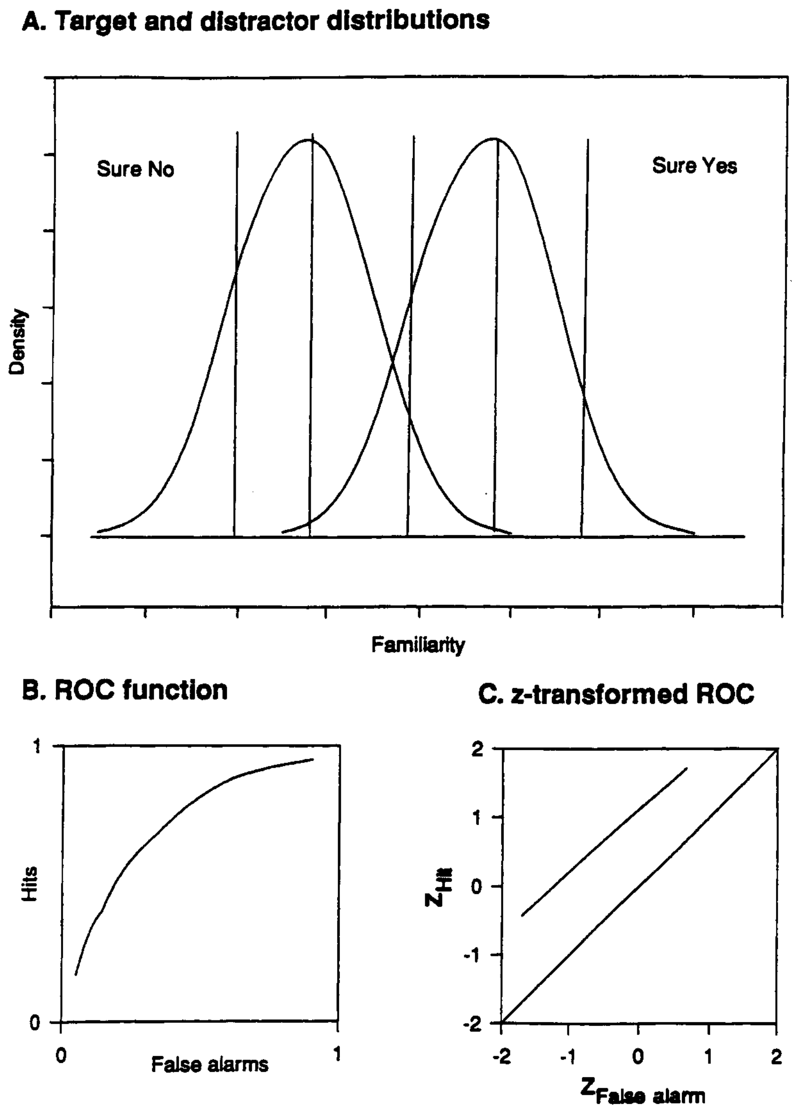

Figure 6. ROC analysis to examine d' and ratio of variances. Panel A shows underlying match and nonmatch distributions with five response criteria that divide the familiarity scale into six confidence regions. Panel B plots the hit rate (vertical axis) and false-alarm rate (horizontal axis) at each response criterion. Panel C shows the z-transformed ROC function, which is linear if the underlying familiarity distributions are normal. The intercept is an estimate of d', and the slope is the ratio of the standard deviations for the nonmatch and match standard deviations.

tersection. The probability of a positive response for a negative item in the exclude condition is

$$\text{P(exclude)} = F - RF,$$

which is the probability that the item is familiar but not recalled. Presumably, one effect of recall is to access information that would identify the item as a negative item. From this, an estimate of the contribution of recall is given by $P(\text{include}) - P(\text{exclude})$, and F is estimated by substituting in R and solving either of the two equations. Yonelinas and Jacoby's analysis indicated that R decreased and F did not change with increases in list length. Thus, their analysis is consistent with the two-factor model proposed earlier by Atkinson and Juola (1974) in that list-length effects are due to recall rather than familiarity.

Yonelinas (1994) also applied this model to ROC analyses. The familiarity-based component of the model assumes equal-variance, normal distributions for all test conditions (only the means vary). Although the processes underlying recall are not specified, the proportional contribution of recall is derived using the process-dissociation method. The process-dissociation estimates of R and F showed that the contribution of recall decreased with increasing list length. As recall's contribution decreases, the slope of the ROC should increase or approach 1.0, which is precisely what Yonelinas' results showed.

Yonelinas found relatively large differences in ROC slopes (.53 vs. .71 for short and long, respectively), whereas Gronlund and Elam found no difference in slopes (approximately 1.0), despite using a greater difference in list lengths. How can we reconcile these results?

One solution might be that the two-list process-dissociation paradigm used by Yonelinas obligates the use of recall when it might not otherwise occur. If the contribution of recall differs between two list lengths, Yonelinas argues that the slopes should differ; however, if the recall contribution does not differ between two list lengths, or if there is no recall contribution, the slopes would not differ. We can think of the Gronlund and Elam (1994) and Yonelinas (1994) results as consistent with the idea that familiarity alone produces slopes unaffected by list length if recall does not contribute (Gronlund & Elam, 1994) or if recall is algebraically removed (Yonelinas, 1994). An experiment that holds list length constant but varies the likelihood of recall could test this; the slope should be less when the likelihood of recall is greater.

It is important to note that the two-process model is a framework for estimating the proportional contribution of two component processes in recognition but that it does not specify the processes themselves. It remains to be seen whether a fully developed process model consistent with this framework would be successful. We will have more to say about this model and the process-dissociation (include–exclude) procedure below.

List-Strength Effect

The same mechanism that produces list-length effects also predicts a list-strength effect, defined as follows: As some items on a list are strengthened, recognition of the other, nonstrengthened items is predicted to decrease. Strengthening here is defined in terms of the duration of a single presentation or through repetition.

The list-strength prediction has been tested using a mixed-list versus pure-list procedure in which subjects study three kinds of lists: A pure-strong list consists of all strong items, a pure-weak list consists of all weak items, and a mixed list consists of half strong and half weak items. According to the models, the list-strength pattern should show that strong items are better remembered when they are presented in a mixed list, and weak items are better remembered when they are presented in a pure-weak list. Put another way, strong items are remembered better and weak items are remembered more poorly in a mixed list than they are in a pure list.

The prediction is illustrated by considering a test item, T, and a different item in memory, item N. As the strength of N increases, the variance of the match between T and N increases. Therefore, as items in a list are strengthened, the variance of the global match increases (the variance of the global match is the sum of the variances of the individual matches, assuming the items are independent). As the variance of the global match increases (with the mean difference between target and distractor distributions unchanged), the distributions for targets and distractors will overlap more and recognition performance will decrease.

The list-strength prediction is due to the ordering of the variances. The variance is largest for the pure-strong list because all of the items are strong, smallest for the pure-weak list because all of the items are weak, and intermediate for the mixed list: thus, var(strong) > var(mixed) > var(weak). Because the expected match (the mean familiarity) for strong items is the same in the pure-strong and mixed lists, but the variance is smaller in the mixed list, performance for strong items should be better in the mixed list. By the same reasoning, the expected match for weak items is the same in the pure-weak list and in the mixed list, but the variance is larger in the mixed list than in the pure-weak list. Consequently, performance should be worse for weak items in the mixed list. The combination of these two factors can be summarized by noting that the difference in performance between strong and weak items should be larger in a mixed-list comparison than in a pure-list comparison.

This lengthy discussion of the origin of the list-strength effect is followed by a rather brief description of the empirical results: It occurs for recall, but several experiments have failed to find a list-strength effect in item or associative recognition (Hockley, 1992b; Murnane & Shiffrin, 1991b; Ratcliff et al., 1990; Yonelinas, Hockley, & Murdock, 1992). In fact, in several experiments, negative list-strength results were found in which the difference in performance for strong and weak items was larger in the pure list than in the mixed list.

The combination of these three results—a list-length effect for recall and recognition, a list-strength effect for recall, and an absent or negative list-strength effect for recognition—provides a difficult problem for global matching models. A solution to the problem would seem to require a decoupling of mean and variance assumptions and a decoupling of processes underlying recall and recognition (because list-strength effects do occur for recall but not for recognition). Two solutions have been offered: a differentiation hypothesis was proposed within the framework of the SAM model (Shiffrin et al., 1990), and a continuous memory hypothesis was proposed within the framework of TODAM (Murdock & Kahana, 1993). Each of these is discussed below.

The differentiation hypothesis. The differentiation hypothesis was developed within the framework of the SAM model. In the standard version of the model, as a given item is studied longer, the cue-to-image strengths will increase between that image and the items with which it is rehearsed. What about the items with which the strengthened item is not rehearsed? In the standard version of the model, these cue-to-image strengths are set to a *constant* residual value (d) that does not vary with study time.

The differentiation hypothesis assumes that the residual strengths do not remain constant with increases in study time but, in fact, decrease. Thus, as a given item is studied longer, it will become more connected to context and to items with which it is rehearsed (standard assumption) and less connected to items with which it was not rehearsed. The second part of this statement is the heart of the differentiation hypothesis: with increased study time, an item becomes differentiated from other items in the list.

The list-strength prediction is produced by the fact that the variance increases with the mean. However, according to the differentiation hypothesis, as the strength of an item increases, residual strengths decrease and, thus, the variance will stay the same or even decrease. Thus, the differentiation mechanism allows recognition performance to increase with study time, without concomitant increases to the variance. The differentiation mechanism allows for a great deal of flexibility, allowing positive, negative, and missing list-strength effects. However, the exact differentiation function and the parameters that govern it have not yet been spelled out. Also, because there is no repetition of items in a list-length experiment, the differentiation assumption does not apply. As a result, SAM with differentiation still incorrectly predicts that variance ratios converge to 1.0 with increasing list length.

Continuous memory assumption. The list-strength prediction is produced by differences in the variances among pure-strong, pure-weak, and mixed lists. The list-strength prediction, therefore, assumes that retrieval is focused on a single list. Murdock and Kahana (1993) have challenged this assumption, proposing a continuous memory that accumulates not only over lists but over preexperimental experiences as well. Thus, variance differences due to list composition disappear, and the list-strength prediction for recognition disappears with it. Murdock and Kahana implemented the continuous memory assumption within the framework of TODAM, by assuming that (1) the memory vector is not initialized to zero prior to list presentation but rather is filled with values representing information in memory prior to the experiment, and (2) over the course of the experiment, the memory vector is not "emptied" after each list—rather, information accumulates continuously in memory.

Although the continuous memory solution is logically sound, it still has some problems. First, not only does it eliminate the variance component underlying the list-strength effect, it also eliminates the variance component underlying the list-length effects. This is acceptable for TODAM, since it can produce list-length effects by mechanisms other than increases in noise (i.e., $\alpha < 1$), but would be a problem for the other models. However, it raises another problem. TODAM would produce list-strength effects by *not* averaging across serial positions, and list-length effects by averaging across serial positions. If TODAM is constrained to predict both effects in the same way, it can predict either the absence of both effects or the presence of both (Shiffrin, Ratcliff, Murnane, & Nobel, 1993).

A second implication of the continuous memory assumption is that it eliminates the list-strength effect for recall. Thus, some other mechanism must be identified as underlying the list-strength effect in recall. One possibility is output interference: strong items are recalled first and interfere with the subsequent recall of the weaker items. However, this mechanism does not account for the list-strength effect shown in cued recall, because the test trials are not ordered according to strength.[10]

A third problem is that the continuous memory assumption cannot predict a negative list-strength effect. The negative list-strength effects shown in the data must be produced by some other mechanism. One possibility is that they are produced by rehearsal differences; specifically, negative list-strength effects could be produced by rehearsal redistribution, whereby subjects "borrow" some of the rehearsal time from strong items and use it to study weak items in a mixed list. However, Murnane and Shiffrin (1991a) have shown that negative list-strength effects are not produced by rehearsal redistribution.

In closing our discussion concerning the continuous memory assumption, we wish to emphasize that despite the problems we have outlined, some form of the assumption seems obligatory in any model of memory in the face of a large body of research indicating the role of prior knowledge in the acquisition of new information (e.g., Bartlett, 1932; Bransford & Johnson, 1972; Brewer & Treyens, 1981; Schank & Abelson, 1977).

More on the Variances

It is clear that many important predictions are rooted in the variances of the underlying match distributions. Relevant here is the relationship between the mean and the variance of a given match distribution. Specifically, does the variance of the match distribution increase with the mean? A recent study used an ROC analysis to examine changes in the variance as a function of the strength of studied items (Ratcliff, McKoon, & Tindall, 1994; Ratcliff, Sheu, & Gronlund, 1992). We call this the *item-strength effect*. Predictions have been derived for SAM, TODAM, and MINERVA 2.

Item-strength effect. SAM and MINERVA 2 predict that the slope of the transformed-ROC function will decrease as the strength of the studied items increases, because the variance of the studied items increases with its mean. In TODAM, on the other hand, the slope is constant near 1.0. The details underlying this derivation are spelled out in detail by Ratcliff et al. (1992). Again, the empirical results are straightforward and contrary to the predictions. The slopes remain *constant* with increases in strength (contrary to SAM and MINERVA 2) and are significantly less than 1.0 (contrary to TODAM). This suggests that the variance for studied items is slightly larger than that for nonstudied items; however, items studied for a longer duration (strong items) did not show a larger variance than did items studied for a shorter duration (weak items).

Perhaps the addition of recall to the models can help explain the ROC data. Yonelinas (1994) applied the process-dissociation analysis to item-strength ROC functions. He showed that if an experimental variable (e.g., increased study time) was positively correlated with both F and R, then the ROC slope can remain constant while the intercept changed. This is an intriguing explanation for the item-strength data. However, it is unclear how circumscribed the increases in F and R must be to maintain constant slopes, and it is not clear that any process model can behave in the manner that these R and F estimates suggest.

The results based on these ROC analyses leave us in a theoretical quandary. The variances appear invariant with strength manipulations, irrespective of whether this results from strengthening the item itself or the other items on the list (Ratcliff et al., 1992; Shiffrin et al., 1990). Although the variances do increase with list length, they stay in the same ratio (Gronlund & Elam, 1994), or perhaps any ratio differences are due to the differential contribution of recall (Yonelinas, 1994).

Differentiation can correct SAM's list-strength mispredictions but is mute regarding the list-length and item-strength ROC data. The continuous memory assumption seems like a necessary addition to the models; however, according to Shiffrin et al. (1993), it does little to improve mispredictions regarding the list-strength effect. Nevertheless, a consequence of a continuous memory—increased difficulty distinguishing between lists—may help explain Gronlund and Elam's (1994) Experiment 2 data. Their Experiment 2 was an exact replication of Experiment 1, with the exception that subjects completed only one list, long or short, rather than multiple long and short lists. In Experiment 2, the variance ratio did increase slightly with increasing list length. Perhaps variance differences between lists can be detected only when extralist variance is minimized, as is the case when only one list is studied and tested. One instantiation of a continuous memory tried by Gronlund and Elam was unsuccessful, but other instantiations are possible.

Global Similarity Effects

The global nature of the matching process provides a parsimonious account of what we term *global similarity effects* in which similarity to a test probe is distributed over multiple items stored in memory. In Posner and Keele's (1970) classic experiment, subjects were presented with a set of dot patterns, generated as deviations from a prototype. Although subjects never saw the prototype pattern during study, there was a strong tendency to call the prototype pattern "old." Others (Bransford & Franks, 1971) have shown a similar pattern of results: a high false-alarm rate for distractor items that are similar to many items in memory but identical to none. The classic account of these prototype effects was that a mental representation of the (nonstudied) prototype pattern was abstracted and stored in memory along with the studied patterns.

This result has since been shown to be consistent with a model in which the test item (i.e., the prototype) is matched to multiple memory traces that are summed together at the time of test (Hintzman & Ludlam, 1980; Medin & Schaffer, 1978). A wide range of prototype effects has recently been produced by various global matching models, including MINERVA 2 (Hintzman, 1986), SAM (Clark, 1988), CHARM (Eich, 1982), and a matrix association model similar to Pike's model (Knapp & Anderson, 1984). Prototype effects are explained very simply in global memory models because a sum of many partial matches to several memory traces may equal or exceed an exact match to a single identical trace (see also Nosofsky, 1988).

In an experiment by Hintzman (1988), subjects studied a categorized list with one to five instances per category. As the number of category items increases, the match for targets and distractors in that category should also increase. Hintzman showed this to be the case for MINERVA 2, and it clearly holds for the other global matching models as well. The prediction, which is borne out in the data, is that hit rates and false-alarm rates both increase as the number of related items on the list increases.

Hintzman, Curran, and Oppy (1992) examined a related prediction, which they termed *proportionality*: as item A is repeated in a list, the judgment of frequency (JOF) for a similar distractor item A' will increase proportionally. Their results, however, showed a more complicated bimodal pattern: the JOF for A' did increase proportionally to the frequency of A, but there was also a large number of zero-frequency estimates as well. Hintzman et al.'s account of this bimodal distribution is straightforward: JOF is based on global matching and therefore increases proportionally, unless one can recall A from the list, which would allow one to rule A' out entirely. This recall-and-reject strategy is quite reasonable in their experiments because similarity between two items was based on plural and singular forms of the same word. Thus, recalling that HORSES was on the list would provide good evidence that HORSE was not.

Hintzman and Curran (1994) found converging evidence for the recall-to-reject hypothesis using a response-signal procedure. Like Gronlund and Ratcliff (1989), they showed a nonmonotonic function for false alarms to similar distractors. As the signal lag was increased, false-alarm rates initially increased for similar distractors but then decreased with additional increases in signal lag.

Recall Processes Revisited

In accord with many other results discussed in this paper, the Yonelinas (1994) account of the item-strength data and the violations of proportionality suggest a role for recall-like retrieval processes in recognition. This seems particularly straightforward with respect to prototypes and proportionality. A distractor that corresponds to a prototype will have high false-alarm rates because it partially matches several items to which it is similar. Thus, the similarity of the prototype is distributed across several items in the list. This makes the prototype difficult to reject for at least two reasons. First, if recall is similarity-based, then none of the items in memory may be similar enough to the prototype to be recalled; second, even if an item is retrieved, it may not serve as evidence to reject the prototype. In Hintzman et al.'s (1992) experiments, the distractor was highly similar to only one item in memory (which may have multiple representations), so the basis of the similarity was such that it was clear that retrieval of a plural item from the list was grounds for rejection of the singular form as a distractor.

CURRENT STATUS OF GLOBAL MEMORY MODELS

Global memory models have provided an important theoretical framework for recognition memory. Our goals in this paper were (1) to trace the origins of the models in relation to previous models; (2) to show how the models are related and how they differ, both conceptually and with respect to data; (3) to work through each model in detail, showing how each model's theoretical machinery leads to its predictions; and (4) to evaluate the models against current data.

The assumption that recognition is based on global matching solves an important problem; recognition decisions are based on global characteristics of memory (i.e., memory of events other than the test items), and yet decisions can in principle be made quickly and accurately. Global matching replaced two-process accounts that combined a fast, direct-access component with a slower recall-like component.

The interactive combination of cues accounts for context effects (Clark & Shiffrin, 1992; Light & Carter-Sobell, 1970; Tulving & Thomson, 1971), and the global matching of test items to memory can account for prototype, list-length, and proportionality effects. However, as some problems were solved, new problems emerged. To some degree, the new problems are related to the old ones: The mechanism that produces list-length effects incorrectly predicts list-strength effects and cannot explain constant variance ratios as a function of list length. Associative recognition and context effects, once thought to reflect the operation of a recall process (Humphreys, 1978; Mandler, 1980), again appear to require the assumption of recall, although on the basis of somewhat different data (Clark, 1992; Clark, Hori, & Callan, 1993). Violations in proportionality are also explainable in terms of a recall process (Hintzman et al., 1992). However, some of these results, while they are suggestive of, or consistent with, the operation of recall processes in recognition, do not provide conclusive evidence for it. The alternative to dissociation at the level of process is, of course, dissociation at the level of information.

Process Versus Information Dissociation

Earlier we described several experiments that show different patterns of results for item and associative recognition. Moreover, associative recognition shows a pattern similar to that for recall with respect to word frequency (Clark, 1992), imagery (McGee, 1980), and encoding instructions (Greene, 1990). In addition, the nonmonotonic retrieval dynamics may indicate recall-like retrieval processes (Gronlund & Ratcliff, 1989). However, although these results can be interpreted in terms of a process dissociation (familiarity vs. recall), they may also be interpreted in terms of an information dissociation. Specifically, we argued that associative information may be represented as a higher order unit that is independent of its item-level components.

Process dissociations may operate in other ways as well. Response-signal experiments by Dosher (1984) and Dosher and Rosedale (1991) required subjects to distinguish between studied pairs and pairs consisting of semantically related items. Their results showed the nonmonotonic false-alarm rate pattern in which false-alarm rates for semantically related but nonstudied lures initially increased but then were suppressed as signal lag increased. This, of course, is the pattern that has been typically taken as evidence for a slow recall process, but Dosher and Rosedale suggest that it may also reflect retrieval of contextual information that connects studied word pairs to the specific episodic context of the experiment. This contextual information may be retrieved by recall-like processes; however, such an assumption is not necessary. Likewise, the fit of the SAM model by Ratcliff, Van Zandt, and McKoon (1995) to Yonelinas's results, which were taken as evidence for recall, was also produced through differential weighting of contextual cues.

Information dissociations may account for various other kinds of task dissociations that have been taken as evidence for multiple processes or multiple systems. Dissociations between implicit and explicit memory tasks (e.g., word-fragment completion and recognition memory) have been taken as evidence of multiple memory systems (Schacter & Tulving, 1994). However, implicit–explicit memory differences may reflect a reliance on different kinds of information in implicit and explicit tasks. For example, explicit memory tasks may incorporate contextual cues into the memory probe to restrict access to the experimental context, whereas implicit tasks (with the same nominal test item) do not (Humphreys, Bain, & Pike, 1989).

Role of Recall in a Two-Process Model

It is important to note that when one speaks of recall processes in recognition, one need not, and should not, imply that all of the processes that underlie recall are exported in whole and inserted into recognition. Recall consists of a number of subprocesses, some of which may operate in recognition, but others of which may not. There are several implications of this. First, one should not estimate the contribution of recall in recognition from performance in recall tasks. Estimating recall in recognition from cued-recall data assumes that recall processes operate in exactly the same way in both tasks, which seems unlikely. Related to this, it also seems unlikely that recall should work in precisely the same way for include and exclude discriminations. In the include condition, recall provides redundant information, whereas in the exclude condition, it is the crucial information that must override familiarity in order to give a correct response.

Recall to Reject

Many of the proposals for recall processes in recognition are of a similar form—that a distractor item can be used as a cue to recall a target item, the presence of which is taken as evidence by the subject that the distractor cannot therefore be from the list. Such proposals have been put forth by Clark (1992) in accounting for word-frequency effects, by Gronlund and Ratcliff (1989) in accounting for the nonmonotonic retrieval dynamics of associative recognition, and by Hintzman et al. (1992) in accounting for violations of proportionality.

The recall-to-reject hypothesis is central to the include–exclude comparison, which defines the process-dissociation procedure. This procedure has become widely used of late and is taken by many to provide strong evidence for the necessity of a recall process in recognition. Thus, it is important to consider this procedure carefully.

The logic underlying the inclusion/exclusion paradigm is that items that are well recognized (included) are also well rejected (excluded). Thus, items in set A will have higher hit rates when the task is to respond positively to them, and they will have low false-alarm rates when the task is to respond negatively to them. These results are difficult to explain within a single-process familiarity model, according to which the high hit rate is presumably due to high familiarity, which should in turn lead to higher, not lower, false positives in the exclude condition. The difference between the include and exclude conditions (for positive responses) provides an estimate of the contribution of recall. Thus, the recall-to-reject postulate is a central component of the process-dissociation logic.

However, include–exclude differences may not always reflect recall. Yonelinas's (1994) subjects studied two lists of words. In one condition, they were to respond positively to List 1 items (include) and negatively to List 2 items (exclude) and nonstudied items. In another condition, the list-to-response mapping was reversed (new items were always excluded). The results showed an include–exclude difference of .25 to .43. However, Ratcliff et al. (1995) were able to fit the data with the (single-process) SAM model. How can a single-process model fit data that are presumably such solid evidence for the operation of two processes? Yonelinas's task required subjects to focus retrieval on a particular list. In SAM, this is accomplished by differential weighting of List 1 and List 2 context cues. This was precisely the approach taken by Ratcliff et al. (1995) to mimic the data.

Role of Global Matching Necessary in a Two-Process Model

If we add a recall component, does the assumption of global matching become superfluous? Why not revert back to a local-match-plus-recall model? Jacoby's two-process model can be viewed as just such a model. According to Yonelinas and Jacoby's (1994) analyses, the familiarity component plays no role in list-length effects, and the recall component in tandem with the assumption of equal variance for targets and distractors is able to fit ROC functions (Yonelinas, 1994). Moreover, a local match model is not bound to the (incorrect) prediction of list-strength effects.

It might seem that the two-process model went by the wayside not because it failed to account for data but rather because it was less parsimonious than the single-process matching mechanism. Are there results that *require* the global matching assumption, results that cannot be accounted for by a local-match-plus-recall model? The question may be rephrased: If we assume a familiarity-plus-recall model, is there evidence that the familiarity is due to a global match rather than a local match? Assuming that list-length effects are due to recall, and given that list-strength predictions are not shown in recognition, what evidence is there for global matching?

Global similarity effects may turn out to provide the strongest evidence for global matching. In particular, Hintzman's (1988) demonstration that false-alarm rates increase with the number of related items in a studied list cannot be accounted for either by local matching (because the similarity is distributed over multiple items) or by recall (which would provide a basis for rejecting the distractor). Experiments of this kind may provide an important playing field on which to play out the next round between global matching and two-process models.

Conclusions and Future Directions

In the introduction of this paper, we discussed the global matching approach as a potential solution to the problem of how recognition decisions can be made quickly and yet be sensitive to multiple items in memory. We think the approach has offered a viable solution to that basic problem (although specific reaction time models still need development).

Moreover, in the last decade, the global matching approach has illuminated a wide range of new questions that address fundamental questions about forgetting, interference, and the nature of associations. Such questions typically have had complicated answers. Global matching models have been extremely useful, not only in making sense of complicated results but also in generating and refining new questions. From the global matching approach, questions about interference and forgetting were taken on within an investigation of the relationship between strengthening versus adding information to memory (list strength vs. list length). Questions about the nature of associations were addressed in terms of higher order associations, the relationship between item and associative information, and the combination of multiple cues.

Many of the results we have reviewed pose serious challenges for the models, and it seems that some of their core assumptions need to be revised. Nonetheless, we anticipate that these models will continue to drive theoretical development in recognition memory, as the discovered inadequacies of the models set the stage for future research and theoretical development.

REFERENCES

ANDERSON, J. R. (1972). FRAN: A simulation model of free recall. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 5, pp. 315-378). New York: Academic Press.

ANDERSON, J. R., & BOWER, G. H. (1973). *Human associative memory*. Hillsdale, NJ: Erlbaum.

ATKINSON, R. C., & JUOLA, J. F. (1974). Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology: Vol. 1. Learning, memory, and thinking* (pp. 243-293). San Francisco: W. H. Freeman.

BAIN, J. D., & HUMPHREYS, M. S. (1988). The relational context effect: Cues, meanings or configurations? In G. Davies & D. M. Thomson (Eds.), *Memory in context: Context in memory* (pp. 97-137). Chichester, U.K.: Wiley.

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

BJORK, R. A., & RICHARDSON-KLAVEHN, A. (1989). On the puzzling relationship between environmental context and human memory. In C. Izawa (Ed.), *Current issues in cognitive processes: The Tulane Flowerree Symposium on Cognition* (pp. 313-344). Hillsdale, NJ: Erlbaum.

BOWER, G. H. (1972). Mental imagery and associative learning. In L. W. Gregg (Ed.), *Cognition in learning and memory* (pp. 51-88). New York: Wiley.

BOWLES, N. L., & GLANZER, M. (1983). An analysis of interference in recognition memory. *Memory & Cognition*, **11**, 307-315.

BRANSFORD, J. D., & FRANKS, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology*, **2**, 331-350.

BRANSFORD, J. D., & JOHNSON, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning & Verbal Behavior*, **11**, 717-726.

BREWER, W. F., & TREYENS, J. C. (1981). Role of schemata in memory for places. *Cognitive Psychology*, **13**, 207-230.

CLARK, S. E. (1988). *A model for classification and memory retrieval*. Unpublished doctoral dissertation, Department of Psychology, Indiana University, Bloomington.

CLARK, S. E. (1992). Word frequency effects in associative and item recognition. *Memory & Cognition*, **20**, 231-243.

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

CLARK, S. E., & BURCHETT, R. E. R. (1994). Word frequency and list composition effects in associative recognition and recall. *Memory & Cognition*, **22**, 55-62.

CLARK, S. E., & HORI, A. (1995). List length and overlap effects in forced-choice associative recognition. *Memory & Cognition*, **23**, 456-461.

CLARK, S. E., HORI, A., & CALLAN, D. E. (1993). Forced-choice associative recognition: Implications for global-memory models. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 871-881.

CLARK, S. E., & SHIFFRIN, R. M. (1987). Recognition of multiple-item probes. *Memory & Cognition*, **15**, 367-378.

CLARK, S. E., & SHIFFRIN, R. M. (1992). Cuing effects and associative information in recognition memory. *Memory & Cognition*, **20**, 580-598.

DAPOLITO, F., BARKER, D., & WILANT, J. (1972). The effects of contextual changes on component recognition. *American Journal of Psychology*, **85**, 431-440.

DOSHER, B. A. (1984). Discriminating preexperimental (semantic) from learned (episodic) associations: A speed-accuracy study. *Cognitive Psychology*, **16**, 519-555.

DOSHER, B. A., & ROSEDALE, G. (1991). Judgments of semantic and episodic relatedness: Common time-course and failure of segregation. *Journal of Memory & Language*, **30**, 125-160.

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, & Cognition*, **16**, 813-824.

EICH, J. M. (1982). A composite holographic associative recall model. *Psychological Review*, **89**, 627-661.

FERNANDEZ, A., & GLENBERG, A. M. (1985). Changing environmental context does not reliably affect memory. *Memory & Cognition*, **13**, 333-345.

GILLUND, G., & SHIFFRIN, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, **91**, 1-65.

GLUCKSBERG, S., & MCCLOSKEY, M. (1981). Decisions about ignorance: Knowing that you don't know. *Journal of Experimental Psychology: Human Learning & Memory*, **7**, 311-325.

GREENE, R. L. (1990). Memory for pair frequency. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 110-116.

GREGG, V. H. (1976). Word frequency, recognition and recall. In J. Brown (Ed.), *Recall and recognition* (pp. 183-216). Chichester, U.K.: Wiley.

GRONLUND, S. D. (1986). *Multi-level storage and retrieval: An empirical investigation and theoretical analysis*. Unpublished doctoral dissertation, Department of Psychology, Indiana University, Bloomington.

GRONLUND, S. D., & ELAM, L. E. (1994). List-length effect: Recognition accuracy and the variance of underlying distributions. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 1355-1369.

GRONLUND, S. D., & RATCLIFF, R. (1989). Time course of item and associative information: Implications for global memory models. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 846-858.

GRONLUND, S. D., & RATCLIFF, R. (1991). Analysis of the Hockley and Murdock decision model. *Journal of Mathematical Psychology*, **35**, 319-344.

GRONLUND, S. D., SHEU, C.-F., & RATCLIFF, R. (1990). Implementation of global memory models with software that does symbolic computation. *Behavior Research Methods, Instruments, & Computers*, **22**, 228-235.

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

HASHER, L., & ZACKS, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, **108**, 356-388.

HINTZMAN, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers*, **16**, 96-101.

HINTZMAN, D. L. (1986). Schema abstraction in a multiple-trace memory model. *Psychological Review*, **93**, 429-445.

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

HINTZMAN, D. L., & CURRAN, T. (1994). Retrieval dynamics of recognition and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory & Language*, **33**, 1-18.

HINTZMAN, D. L., CURRAN, T., & OPPY, B. (1992). Effects of similarity and repetition on memory: Registration without learning? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 667-680.

HINTZMAN, D. L., & LUDLAM, G. (1980). Differential forgetting of prototypes and old instances: Simulation by an exemplar-based classification model. *Memory & Cognition*, **8**, 378-382.

HINTZMAN, D. L., & WATERS, R. M. (1970). Recency and frequency as factors in list discrimination. *Journal of Verbal Learning & Verbal Behavior*, **9**, 218-221.

HOCKLEY, W. E. (1991). Recognition memory for item and associative information: A comparison of forgetting rates. In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock* (pp. 227-248). Hillsdale, NJ: Erlbaum.

HOCKLEY, W. E. (1992a). Item versus associative information: Further comparisons of forgetting rates. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 1321-1330.

HOCKLEY, W. E. (1992b, November). *Tests of the mirror and list-strength effects for item and associative recognition.* Paper presented at the meeting of the Psychonomic Society, St. Louis.

HOCKLEY, W. E. (1994). Reflections of the mirror effect for item and associative recognition. *Memory & Cognition*, **22**, 713-722.

HOCKLEY, W. E., & CHRISTI, C. (1996). Tests of the differential encoding of 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.

HOVLAND, C. I., & KURTZ, K. H. (1952). Experimental studies in rote learning theory: X. Prelearning syllable familiarization and the length-difficulty relationship. *Journal of Experimental Psychology*, **44**, 31-39.

HUMPHREYS, M. S. (1976). Relational information and the context effect in recognition memory. *Memory & Cognition*, **4**, 221-232.

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

HUMPHREYS, M. S., & BAIN, J. D. (1983). Recognition memory: A cue and information analysis. *Memory & Cognition*, **11**, 583-600.

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., 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.

HUNT, R. R., & EINSTEIN, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning & Verbal Behavior*, **20**, 497-514.

JACOBY, L. L. (1991). A process dissociation framework: Separating automatic and intentional uses of memory. *Journal of Memory & Language*, **30**, 513-541.

KINTSCH, W. (1970). Models for free recall and recognition. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press.

KNAPP, A. G., & ANDERSON, J. A. (1984). Theory of categorization based on distributed memory storage. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 616-637.

LIGHT, L. L., & CARTER-SOBELL, L. (1970). Effects of changed semantic context on recognition memory. *Journal of Verbal Learning & Verbal Behavior*, **9**, 1-11.

MANDLER G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, **87**, 252-271.

MCGEE, R. (1980). Imagery and recognition memory: The effects of relational organization. *Memory & Cognition*, **8**, 394-399.

MCKOON, G., & RATCLIFF, R. (1979). Priming in episodic and semantic memory. *Journal of Verbal Learning & Verbal Behavior*, **18**, 463-480.

MEDIN, D. L., & SHAFFER, M. M. (1978). A context theory of classification learning. *Psychological Review*, **85**, 207-238.

MENSINK, G.-J. M., & RAAIJMAKERS, J. G. (1989). A model for contextual fluctuation. *Journal of Mathematical Psychology*, **33**, 172-186.

METCALFE, J., & MURDOCK, B. B. (1981). An encoding and retrieval model of single-trial free recall. *Journal of Verbal Learning & Verbal Behavior*, **20**, 161-189.

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

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

MURDOCK, B. B. (1983). A distributed memory model for serial-order information. *Psychological Review*, **90**, 316-338.

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

MURDOCK, B. B., & HOCKLEY, W. E. (1989). Short-term memory for associations. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 24, pp. 71-108). San Diego, CA: Academic Press.

MURDOCK, B. B., & KAHANA, M. J. (1993). Analysis of the list-strength effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 689-697.

MURNANE, K., & PHELPS, M. P. (1993). A global activation approach to the effect of changes in environmental context on recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 882-894.

MURNANE, K., & PHELPS, M. P. (1994). When does a different environmental context make a difference in recognition? A global activation model. *Memory & Cognition*, **22**, 584-590.

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

MURNANE, K., & SHIFFRIN, R. M. (1991b). Word repetitions in sentence recognition. *Memory & Cognition*, **19**, 119-130.

NEELY, J. H., & DURGUNOGLU, A. Y. (1985). Dissociative episodic and semantic priming effects in episodic recognition and lexical decision tasks. *Journal of Memory & Language*, **24**, 466-489.

NOBEL, P. A., & HUBER, D. E. (1993). Modeling forced-choice associative recognition through a hybrid of global recognition and cued recall. In *Proceedings of the 15th Annual Conference of the Cognitive Science Society* (pp. 783-788). Hillsdale, NJ: Erlbaum.

NOSOFSKY, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 700-708.

PIKE, R. (1984). A comparison of convolution and matrix distributed memory systems. *Psychological Review*, **91**, 281-294.

POSNER, M. I., & KEELE, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology*, **83**(2, Pt. 1), 304-308.

POSTMAN, L. (1962). Effects of language habits on the acquisition and retention of verbal associations. *Journal of Experimental Psychology*, **64**, 7-19.

RAAIJMAKERS, J. G. W., & SHIFFRIN, R. M. (1980). SAM: A theory of probabilistic search of associative memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 14, pp. 207-262). New York: Academic Press.

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

RATCLIFF, R., CLARK, S., & SHIFFRIN, R. M. (1990). List-strength effect: I. Data and discussion. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 163-178.

RATCLIFF, R., MCKOON, G., & TINDALL, M. (1994). Empirical generality of data from recognition memory receiver-operating characteristic functions and implications for the global memory models. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 763-785.

RATCLIFF, R., SHEU, C.-F., & GRONLUND, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review*, **99**, 518-535.

RATCLIFF, R., VAN ZANDT, T., & MCKOON, G. (1995). Process dissociation, single-process theories, and recognition memory. *Journal of Experimental Psychology: General*, **124**, 352-374.

REED, A. V. (1973, August 10). Speed-accuracy trade-off in recognition memory. *Science*, **181**, 574-576.

ROBBINS, D., BRAY, J. F., & IRVIN, J. R. (1974). Intralist contrast effects in cued recall. *Journal of Experimental Psychology*, **103**, 150-155.

SCHACTER, D. L., & TULVING, E. (EDS.) (1994). *Memory Systems 1994*. Cambridge, MA: MIT Press.

SCHANK, R. C., & ABELSON, R. (1977). *Scripts, plans, goals, and understanding*. Hillsdale, NJ: Erlbaum.

SHEU, C.-F. (1992). A note on the multiple-trace memory model without simulation. *Journal of Mathematical Psychology*, **36**, 592-597.

SHIFFRIN, R. M., MURNANE, K., GRONLUND, S. D., & ROTH, M. (1988). On units of storage and retrieval. In C. Izawa (Ed.), *Current issues in cognitive processes: The Tulane Flowerree Symposium on Cognition* (pp. 25-68). Hillsdale, NJ: Erlbaum.

SHIFFRIN, R. M., RATCLIFF, R., & CLARK, S. (1990). List-strength effect: II. Theoretical mechanisms. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 179-185.

SHIFFRIN, R. M., RATCLIFF, R., MURNANE, K., & NOBEL, P. (1993). TODAM and the list-strength and list-length effects: Comment on Murdock and Kahana (1993). *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 1445-1449.

SMITH, S. M. (1988). Environmental context-dependent memory. In G. M. Davies & D. M. Thomson (Eds.), *Memory in context: Context in memory*. Chichester, U.K.: Wiley.

SMITH, S. M., GLENBERG, A., & BJORK, R. A. (1978). Environmental context and human memory. *Memory & Cognition*, **6**, 342-353.

STRONG, E. K., JR. (1912). The effect of length of series upon recognition memory. *Psychological Review*, **19**, 447-462.

THOMSON, D. M. (1972). Context effect in recognition memory. *Journal of Verbal Learning & Verbal Behavior*, **11**, 789-793.

THOMSON, D. M., & TULVING, E. (1970). Associative encoding and retrieval: Weak and strong cues. *Journal of Experimental Psychology*, **86**, 255-262.

TULVING, E. (1976). Ecphoric processes in recall and recognition. In J. Brown (Ed.), *Recall and recognition*. London: Wiley.

TULVING, E. (1981). Similarity relations in recognition. *Journal of Verbal Learning & Verbal Behavior*, **20**, 479-496.

TULVING, E., & SCHACTER, D. L. (1990, January 19). Priming and human memory systems. *Science*, **247**, 301-306.

TULVING, E., & THOMSON, D. M. (1971). Retrieval processes in recognition memory: Effects of associative context. *Journal of Experimental Psychology*, **87**, 175-184.

UNDERWOOD, B. J., RUNQUIST, W. N., & SCHULZ, R. W. (1959). Response learning in paired-associate lists as a function of intralist similarity. *Journal of Experimental Psychology*, **58**, 70-78.

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

YONELINAS, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 1341-1354.

YONELINAS, A. P., HOCKLEY, W. E., & MURDOCK, B. B. (1992). Tests of the list-strength effect in recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 345-355.

YONELINAS, A. P., & JACOBY, L. L. (1994). Dissociations of processes in recognition memory: Effects of interference and response speed. *Canadian Journal of Experimental Psychology*, **48**, 516-534.

YUILLE, J. C., & PAIVIO, A. (1967). Latency of imaginal and verbal mediators as a function of stimulus and response concreteness-imagery. *Journal of Experimental Psychology*, **75**, 540-544.

NOTES

1. A somewhat more technical review of the models is given by Humphreys, Pike, Bain, and Tehan (1989).

2. TODAM and CHARM may be viewed as variants of a single model, sharing similar representational and retrieval assumptions. Therefore, discussion of both models would be somewhat redundant. We have elected to focus on TODAM rather than CHARM, because TODAM has been applied to the results and issues that are relevant to this article to a greater extent than has CHARM. CHARM has provided substantial contributions in several areas, notably similarity and auto-associations.

3. This simplicity allows the predictions of the models to be easily obtained. Equations for the means and variances can be derived in many instances and are listed for various test conditions in various sources (for TODAM, see Weber, 1988; for the Matrix model, see Pike, 1984; for MINERVA 2, see Sheu, 1992; for SAM, see Clark & Shiffrin, 1992. Clark & Shiffrin provide equations for all of the global matching models for several test conditions, and Gronlund, Sheu, & Ratcliff, 1990, provide software for doing derivations for TODAM and SAM). In cases where the means and variances cannot be obtained by derivation, predictions can be obtained by computer simulation. Simulation programs are often available from the original authors.

4. We organize the paper with these two major assumptions in order to emphasize comparisons that transcend particular models. It should be noted, however, that theoeretical assumptions must be evaluated within the context of the model in which they are implemented. For this reason, and because the same assumption may be implemented in a variety of ways, assumptions cannot be tested in a "divide-and-conquer" fashion (borrowing a metaphor from Ben Murdock).

5. The recall-augmented model proposed by Humphreys is different from those of Atkinson and Juola (1974) and Mandler (1980). In Humphreys's (1978) model, recall comes into play only when multiple items are tested. Because Gillund and Shiffrin (1984) tested with single items, their results do not provide a direct test of Humphreys's model.

6. The "in principle" qualification is added because a *detailed* reaction time model has been developed only for TODAM (Hockley & Murdock, 1987), and this model has been shown to have difficulties with certain details of reaction time data (see Gronlund & Ratcliff, 1991). Nonetheless, the parallel nature of the models would allow for fast reaction times.

7. The approximate familiarity values do not take into account the cue weights w_i in Equation 1 or assumptions concerning the variance of cue-to-image strengths. The variance is implemented by multiplying each retrieval strength by a value υ, which has typically been set to take on values .5, 1.0, and 1.5 with equal probability.

8. One might think that the solution to the problem is to relax this assumption, allowing $\sum w_i > 1$. However, this has negative side effects, particularly for recall (see Gronlund & Shiffrin, 1986).

9. Humphreys (personal communication, April, 1994) pointed out that the independence between item and associative information in TODAM is the result of zero-centered vectors. If these were not used in TODAM, autoassociation would be required to produce independence between item and associative information.

10. In addition, Robbins, Bray, and Irvin (1974) did manipulate test order in mixed-strength lists and showed no effects of output interference; weak items were not more poorly recalled. Their results are difficult to interpret, however, because the pure-list control lists were matched to the mixed lists based on the total number of presentations, rather than the number of unique items presented. Thus, their list-length and list-strength effects were confounded.

(Manuscript received February 2, 1994;
revision accepted for publication July 18, 1995.)