

# Associative symmetry and memory theory

MICHAEL J. KAHANA

*Brandeis University, Waltham, Massachusetts*

This article reexamines the theory and data concerning two opposing views of episodic association. The independent association hypothesis (IAH) sees associations as unidirectional and separately modifiable links between individual item representations. The associative symmetry hypothesis (ASH) sees an association as a holistic conjunction of the constituent items, blending the elements of their representations into a new one. In early tests of these hypotheses, experimenters compared forward and backward recall of paired associates, looking for asymmetries. Although some studies showed significant differences between forward and backward recall, the vast majority did not. The author of this study used a mathematical analysis of distributed memory models to reexamine this classic question. These analyses revealed that symmetric and asymmetric models can mimic each other, offering identical predictions regarding forward and backward recall. To distinguish these models, the author examined the correlation between forward and backward recall at the level of individual pairs of items. Both this correlation and the correlation between recall of pairs tested in the same direction were near unity. These results provide new evidence favoring the ASH over the IAH.

This article examines a basic question concerning the nature of episodic associations. On the one hand, associations can be viewed as directional pointers connecting different mental representations. While studying a pair of words in a memory experiment (**A**–**B**), subjects strengthen a pointer going from **A** to **B** and a separate one going from **B** to **A**, thus forming two distinct unidirectional associations. This view maps onto the classic independent association hypothesis (IAH), which viewed forward and backward associations as distinct, separately modifiable (e.g., Ebbinghaus, 1885/1913; Robinson, 1932), and even statistically independent (Wolford, 1971).

Arguing against the IAH, Gestalt theorists promoted a principle of associative symmetry (Asch & Ebenholtz, 1962; Köhler, 1947). This alternate view saw an episodically formed association as a holistic conjunction of the **A** and **B** representations, without any directional **A**→**B** or **B**→**A** pointers. According to this associative symmetry hypothesis (ASH), each member of the associated pair can recover the entire pattern, independent of the order of presentation (Asch, 1968; Asch & Ebenholtz, 1962; Rock & Ceraso, 1964). Experimental evidence on the equality of forward and backward associative recall, discussed below, and the arguments in favor of one-trial associative learn-

ing (e.g., Estes, 1960; Rock, 1957) lent support to this position.

With the ascent of computational and distributed memory models, this basic question concerning the nature of episodic associations gained renewed importance. Distributed memory models make explicit assumptions about the mechanisms of association. As we will show in this article, some models assume symmetrical associative learning of item representations, whereas other models allow for separate formation of forward and backward associations.

Studies aimed at discriminating the IAH from the ASH examined the dependence of recall on the temporal order of encoding. This was examined by using the classic paired-associate method. Subjects studied randomly paired meaningful items, denoted **A**–**B**, in temporal succession. Then the experimenter tested each pair by either probing with **A** for recall of **B** (a forward test) or probing with **B** for recall of **A** (a backward test). A finding of asymmetric retrieval (e.g., better forward than backward recall) was seen as evidence against ASH and in favor of IAH. Similarly, a failure to find the expected asymmetry result was taken as support for the ASH, at least on grounds of parsimony (e.g., Asch & Ebenholtz, 1962; Murdock, 1962).

Even if forward and backward recall are equivalent on average, this does not strictly imply symmetry. Rather, the ASH requires that for every studied pair, **A**–**B**, the probability of recalling **B** given **A** perfectly predicts the probability of recalling **A** given **B**. That is, the correlation between forward and backward recall, at the level of individual pairs, must be 1.0. At the other extreme, the IAH (e.g., Wolford, 1971) implies a correlation of 0.0 between forward and backward recall.

Although there have been numerous studies of recall direction in paired associates, as will be reviewed here, none

---

This research was funded by National Institutes of Health Grant MH55687 and by APA Grant 146 from the National Science and Engineering Research Council of Canada. The author is especially grateful to Dan Rizzuto, Jeremy Caplan, and Marc Howard for helpful comments and to Ian Serotkin for programming and running the experiment. Correspondence concerning this article should be addressed to M. J. Kahana, Center for Complex Systems, Brandeis University, Waltham, MA 02454-9110 (e-mail: kahana@brandeis.edu).

—Accepted by previous editorial team

of these studies attempted to measure the correlation between forward and backward recall at the level of individual pairs. Determining this correlation requires the experimenter to test each pair twice, once in the forward direction and once in the backward direction. As in all correlational studies, the precise level of dependency between the outcomes of such successive tests will partially depend on factors that influence the two tests either independently or in concert. These factors complicate the interpretation of correlations and have led some to question the merits of the successive-testing approach (see Kahana, 2000, for a review).

This article aims to relate theory and data on the question of associative symmetry versus independent associations. Part 1 presents a selected review of the empirical evidence pertaining to symmetric retrieval. Part 2 presents a theoretical analysis of symmetric, asymmetric, and independent associations within the framework of a number of mathematical memory models. This analysis reveals that symmetric models can account for some types of asymmetries and that data on the correlation between forward and backward recall are needed to distinguish among models. Part 3 presents an experimental study, using the method of successive tests, to determine the correlation between forward and backward recall. The experimental results are discussed in terms of the theoretical analyses presented in Part 2.

## PART 1

### The Empirical Evidence—A Brief Review

If the IAH is valid, the order in which a sequence is learned could have a differential effect on forward- and backward-recall probabilities. If the ASH is valid, the order of learning should not differentially impact forward and backward recall. These predictions can be tested by comparing average forward- and backward-recall probabilities under a variety of conditions.

One should note, though, that if the variable strengths of forward and backward associations are independent, it is still possible for the average performance in forward- and backward-recall tests to be indistinguishable. If the averages of the strength distributions for forward and backward associations are the same but individual forward and backward strengths are drawn independently, forward- and backward-recall tests will reveal equivalent performances even though the strengths of the associations are independent.

#### Order of Presentation

The belief in stronger forward than backward associations (embodied in the IAH) neatly fits the stimulus–response (S–R) view of associative learning that prevailed in the first half of the 20th century. This belief remained largely unchallenged until the 1950s and 1960s, when a number of investigators reported equivalent forward and backward recall when the nominal S and R items were both drawn from the same population. For example, Mur-

dock (1956) found equal transfer for forward and backward associations in a paired-associate learning paradigm. Asch and Ebenholtz (1962) argued that studies yielding better forward than backward recall do so because of differential item availability. Because anticipation learning requires subjects to generate the response item when cued with the stimulus item, the response item is learned to a greater extent than the stimulus item. Through learning, the response item becomes more available and, thus, easier to recall. On the basis of this evidence, Asch and Ebenholtz proposed the ASH, arguing that an association is a unitary entity, with no preferred directionality imposed by the order of encoding. This hypothesis represented a radical departure from the then-dominant S–R framework.<sup>1</sup>

Murdock (1962, 1965, 1966) conducted an extensive series of experiments to determine whether order of encoding influences retrieval difficulty. Rather than using anticipation learning of nonsense syllables, Murdock presented subjects with short lists of word pairs for a single study trial. He then gave the subjects a cued-recall test on the studied pairs. In each of these studies, he found nearly identical forward and backward recall across (1) variations in serial position, (2) delay between study and test, and (3) number of presentations (1–3). More recent studies have shown that these results are readily replicated. Mandler, Rabinowitz, and Simon (1981) found no difference between forward and backward recall under standard paired-associate study–test conditions. They also found that in free recall of pairs, subjects were far more likely to recall both members of a pair together than they were to recall individual members. They argued that these results support the notion of coordinate organization of word pairs—a concept similar to the notion of associative symmetry.

During the period of 1956–1966, a great number of other studies examined the question of directionality in paired-associate learning. In a comprehensive review of the literature of this period, Ekstrand (1966) concluded that “it does appear that the difference between forward and backward associative learning has been drastically overestimated, and that if symmetry is not the rule, asymmetry will be very small” (p. 60). Many of the studies surveyed by Ekstrand used anticipation learning and different populations of items for stimulus and response terms—conditions that would later be shown to consistently yield asymmetric retrieval.

Although symmetric retrieval is typically found in paired-associate learning, this result is not true of other tasks. For example, in free recall, after a given list item is recalled, the next item that subjects recall is about twice as likely to be the subsequent list item than it is to be the prior one (Kahana, 1996). This asymmetry effect is extremely robust, appearing in recall of (1) both short and long lists, (2) lists presented auditorily and visually, (3) lists presented at fast and slow rates, (4) lists studied by younger and older subjects, and (5) lists of high- and low-frequency words (Kahana, 1996; Kahana, Howard, Zaromb, & Wingfield, 2002; Ward & Woodward, 2002). It even appears when

a demanding distractor task is interpolated between each list item (Howard & Kahana, 1999). In a companion article, Kahana and Caplan (2002) reported striking asymmetries in probed recall of serial lists. Even in paired-associate learning, some conditions produce asymmetries in retrieval. The next three subsections will review conditions that reliably produce asymmetries in recall.

### Anticipation Learning and Response Availability

Following the preliminary results of Asch and Ebenholtz (1962), Horowitz and colleagues found that equality of forward and backward recall was observed only when response availability was equated for both **A** and **B** items (Horowitz, Brown, & Weissbluth, 1964; Horowitz, Norman, & Day, 1966). For example, repeatedly generating the response term, as is done in the anticipation procedure, will enhance the availability of the **B** items over that of the **A** items, thus producing asymmetric recall.

Horowitz et al. (1966) controlled for response availability by having their subjects study double- or triple-function lists (e.g., Primoff, 1938; Slamecka, 1976). For a critical **A–B** pair, **A** served as a response to one or two other items, and **B** served as a stimulus to one or two other items (these are called double- or triple-function lists because each item can serve as a cue for two or three other list items). After a list of such pairs had been studied, Horowitz et al. (1966) asked subjects to give two free associations to each item in the list. Consistent with the ASH, retrieval was symmetric with respect to order of study when both **A** and **B** served a double function (Horowitz et al., 1966, Experiment 2). However, significant asymmetry was observed for terminal pairs in which only one member of the pair served a double function. For single-function paired associates (a control condition), forward recall was easier than backward recall.

Levy and Nevill (1974) found significant asymmetry following learning of alphanumeric pairs, using the anticipation method. Rather than using double-function lists to control for item availability, they argued that alphanumerics were so familiar that repeated generation would not further enhance their availability. This would seem to be strong evidence against the ASH; however, there is reason to doubt whether the use of familiar materials would actually control for item availability. For example, Nairne and Widner (1988) and Gardiner, Gregg, and Hampton (1988) have shown a larger generation advantage for familiar items than for nonfamiliar items. This result suggests that the response items, repeatedly generated in the anticipation procedure, will be easier to recall than the stimulus items, which are read but not generated.

### Drawing A and B From Different Stimulus Classes

Asymmetric recall can be readily observed when the **A** and the **B** items are chosen from different stimulus pools. For example, Lockhart (1969) found asymmetry for adjective–noun pairs and concrete-noun–abstract-noun pairs. In particular, nouns served as better cues than ad-

jectives, and concrete nouns served as better cues than abstract nouns. Order of study, however, had no effect on recall of items. Bartling and Thompson (1977, Experiment 2) replicated Lockhart's results. They found asymmetry for cued recall of adjective–noun pairs, with nouns being the preferred retrieval cue, independent of the order in which subjects studied the word pair. Noun–noun pairs and adjective–adjective pairs, however, exhibited symmetric retrieval.

### Degree of Learning and Response Latencies

Although equating item availability and sampling items from a common pool are reasonable prerequisites for observing symmetric retrieval, several studies meeting these criteria nonetheless detected significant asymmetries. Wollen, Fox, and Lowry (1970) had two groups study word pairs in forward order, but without anticipation. One of the groups was always tested on forward recall, whereas the other group was always tested on backward recall. Study and test trials alternated. Items were presented at a fast rate (one pair per second), and the subjects read words aloud to limit the possibility of covert rehearsal. At intermediate stages of forward learning, the forward-recall group remembered more words than did the backward-recall group.

Waugh (1970) was interested in the effect of overlearning on forward and backward retrieval. Her subjects studied a list of double-function word pairs to a performance criterion. Then she gave the subjects 15 trials of forward anticipations followed by 15 trials of backward anticipations. Response latencies decreased during the first five or six forward anticipations. When the testing switched to backward anticipations, response latencies slowed to near their original level. Waugh interpreted this violation of associative symmetry in terms of the distinction between memory and performance. She suggested that associations may be symmetrical in memory but that, with extensive training, the performance component involved in recalling the pairs acquires distinct directionality.

## PART 2

### Associative Symmetry and Memory Theory

The literature reviewed above dispels the commonly held belief that forward recall is, in general, easier than backward recall. With common words serving as both **A** and **B** items, recall accuracy does not depend on order of study, regardless of variations in presentation rate, serial position, study–test lag, or list length (e.g., Mandler et al., 1981; Murdock, 1962). Nonetheless, asymmetric retrieval can be found under certain conditions. The question is whether these findings distinguish between the ASH and the IAH.

According to the IAH, forward and backward associative strengths are separately modifiable (and independent). Therefore, even if asymmetries in average forward and backward recall were never observed, one still could not reject the IAH. The ASH is far less flexible. It would

seem that any finding of asymmetric retrieval would be fatal for this hypothesis. The appeal of the ASH is that it is more constrained and thus may teach us something important about the psychological processes involved in associative learning, in its success or its failure.

Up to this point, our discussion has been based on informal, verbal arguments. It is not clear, on the basis of these arguments, what exactly is meant by independent associations and associative symmetry. Mathematical memory models can help by providing a framework for evaluating these hypotheses in precise terms. Such analyses, presented below, show how even models that assume symmetric associative processes can account for violations of symmetry.

### Distributed Memory Models

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

An autoassociative mechanism binds features in such a way that a part can be used to retrieve the whole (termed *redintegration*). A heteroassociative mechanism binds features in such a way that one activated pattern can be used to retrieve another. Simplified versions of biological neural networks can be used to achieve both autoassociation and heteroassociation (e.g., McNaughton & Morris, 1987). We focus, however, on abstract mathematical models that are relatives of the more biologically inspired neural network models of autoassociative and heteroassociative memory. One advantage of working with these abstract models is that analytic solutions for their behavior can be obtained more easily than with more biologically realistic, nonlinear neural networks (Weber, 1988).

Two mathematical operations have been proposed to model associative learning. In one approach (Anderson, Silverstein, Ritz, & Jones, 1977; Humphreys, Bain, & Pike, 1989; Humphreys, Pike, Bain, & Tehan, 1989; Pike, 1984), the association is formed by taking the outer product of two  $n$ -dimensional vectors. The result of this operation is an  $n \times n$  matrix. Matrix memory models have enjoyed enormous popularity on account of their analogy to simplified models of neural plasticity and dynamics (T. H. Brown & Chattarji, 1995; McNaughton & Morris, 1987; Treves & Rolls, 1994). In a second approach (D. A. Brown, Dalloz, & Hulme, 1995; Metcalfe, 1985, 1991, 1993; Murdock, 1979, 1982, 1997), the association is formed by taking the vector convolution of two  $n$ -dimensional vectors. The result of this operation is a  $2n - 1$  dimensional vector (see Murdock, 1979, for details). In both cases, if the vectors being associated are identical, the operation is autoassociation. If the vectors being associated are differ-

ent, the operation is heteroassociation. Once these associative representations are formed, they can be added to a single-memory structure. The same structure can represent many different associations.

**Heteroassociative matrix models.** Consider a pair of vectors,  $\mathbf{a}_j$  and  $\mathbf{b}_j$ , representing the  $j$ th pair of items a subject studies in a list. In a heteroassociative matrix model, the storage equation calls for adding the outer product of  $\mathbf{a}_j$  and  $\mathbf{b}_j$  to the memory matrix,  $W$ . That is,

$$W_j = W_{j-1} + \mathbf{b}_j \mathbf{a}_j', \quad (1)$$

where  $j$  is the index of the current to-be-learned pair and the prime denotes the transpose operation (see Jordan, 1986, for a review of the linear algebra used in these models). For a list of  $L$  pairs,

$$W = \sum_{j=1}^L \mathbf{b}_j \mathbf{a}_j'.$$

Multiplying the memory matrix from the right by a cue item,  $\mathbf{a}_k$ , retrieves a linear superposition of all items that are associated with the cue:

$$\begin{aligned} W \mathbf{a}_k &= \sum_{j=1}^L \mathbf{b}_j \mathbf{a}_j' \mathbf{a}_k \\ &= \mathbf{b}_k \mathbf{a}_k' \mathbf{a}_k + \sum_{j \neq k} \mathbf{b}_j \mathbf{a}_j' \mathbf{a}_k \\ &= \mathbf{b}_k (\mathbf{a}_k \cdot \mathbf{a}_k) + \sum_{j \neq k} \mathbf{b}_j (\mathbf{a}_j \cdot \mathbf{a}_k) \\ &= \mathbf{b}_k + \text{error}. \end{aligned} \quad (2)$$

If the item vectors are orthonormal (i.e., uncorrelated and of unit length), retrieval will be perfect (i.e.,  $W \mathbf{a}_k = \mathbf{b}_k$ ). However, with correlated item vectors, the model retrieves a linear combination of items, with the target item being the strongest.<sup>2</sup> In this heteroassociative matrix model, only forward recall is possible. To achieve backward recall, the memory matrix must include both the forward- and the backward-connection matrices (Humphreys, Bain, & Pike, 1989; Pike, 1984).

We can easily augment our heteroassociative matrix model to include both forward and backward associations. Giving variable weights to these separate associative matrices yields the following storage equation:

$$W_j = W_{j-1} + \gamma_f \mathbf{b}_j \mathbf{a}_j' + \gamma_b \mathbf{a}_j \mathbf{b}_j'. \quad (3)$$

The variable weights on the forward and backward associations,  $\gamma_f$  and  $\gamma_b$ , respectively, subsume the many sources of variability in episodic encoding. If the weights on the forward and backward associations are independent ( $\text{cov}[\gamma_f, \gamma_b] = 0$ ) and if forward associations are stronger than backward associations ( $E[\gamma_f] > E[\gamma_b]$ ), this model embodies the IAH. If, instead,  $\gamma_f$  and  $\gamma_b$  are perfectly correlated

$$\text{cov}[\gamma_f, \gamma_b] = \sqrt{\text{var}[\gamma_f] \text{var}[\gamma_b]}$$

and of equal strength ( $E[\gamma_f] = E[\gamma_b]$ ), this model embodies the ASH.

In such a heteroassociative model, one can imagine possibilities intermediate between independent associations and associative symmetry. For example, some factors that influence degree of learning (e.g., global attention) may affect  $\gamma_f$  and  $\gamma_b$  in a correlated fashion, whereas other factors may influence  $\gamma_f$  and  $\gamma_b$  independently. In this case, the correlation between  $\gamma_f$  and  $\gamma_b$  would take on an intermediate value between 0 and 1.

**Autoassociative matrix models.** Whereas heteroassociation is used to store associations, autoassociation is used to create a content addressable memory, storing individual items that can be retrieved when given a partial input (e.g., Hertz, Krogh, & Palmer, 1991; Hopfield, 1982). But there is no reason why autoassociation cannot also be used to store associations between items. To do this, we sum the  $\mathbf{a}$  and the  $\mathbf{b}$  item vectors and then autoassociate the resulting composite vector. Symbolically,

$$\begin{aligned} W_j &= W_{j-1} + (\mathbf{a}_j + \mathbf{b}_j)(\mathbf{a}_j + \mathbf{b}_j)' \\ &= W_{j-1} + \mathbf{a}_j \mathbf{a}_j' + \mathbf{a}_j \mathbf{b}_j' + \mathbf{b}_j \mathbf{a}_j' + \mathbf{b}_j \mathbf{b}_j'. \end{aligned}$$

In this equation, the autoassociation carries item-specific information about  $\mathbf{A}$  and  $\mathbf{B}$ , as well as the relational information needed for both forward and backward recall. Multiplying the cue item,  $\mathbf{a}_k$ , by the memory matrix retrieves an approximate representation of the studied pair,  $\mathbf{a}_k + \mathbf{b}_k$ :

$$\begin{aligned} W \mathbf{a}_k &= \sum_{j=1}^L (\mathbf{a}_j \mathbf{a}_j' + \mathbf{a}_j \mathbf{b}_j' + \mathbf{b}_j \mathbf{a}_j' + \mathbf{b}_j \mathbf{b}_j') \mathbf{a}_k \\ &= (\mathbf{a}_k + \mathbf{b}_k)(\mathbf{a}_k \cdot \mathbf{a}_k + \mathbf{b}_k \cdot \mathbf{a}_k) \\ &\quad + \sum_{j \neq k} (\mathbf{a}_j + \mathbf{b}_j)(\mathbf{a}_j \cdot \mathbf{a}_k + \mathbf{b}_j \cdot \mathbf{a}_k) \\ &= \mathbf{a}_k + \mathbf{b}_k + \text{error}. \end{aligned}$$

Backward retrieval is accomplished in exactly the same way:

$$\begin{aligned} W \mathbf{b}_k &= \sum_{j=1}^L (\mathbf{a}_j \mathbf{a}_j' + \mathbf{a}_j \mathbf{b}_j' + \mathbf{b}_j \mathbf{a}_j' + \mathbf{b}_j \mathbf{b}_j') \mathbf{b}_k \\ &= \mathbf{a}_k + \mathbf{b}_k + \text{error}. \end{aligned}$$

Again, if the item vectors are orthonormal, retrieval will be perfect.

Because vector addition is commutative, this implementation of cued recall is inherently symmetrical. Even if the strengths of the  $\mathbf{a}$  and  $\mathbf{b}$  items are differentially weighted, the forward and backward heteroassociative terms would have the same strength. The symmetric nature of these linear models also applies to their nonlinear variants (e.g., Buhmann, Divko, & Schulten, 1989; Hopfield, 1982).

**Convolution models.** As before, consider a pair of vectors,  $\mathbf{a}_j$  and  $\mathbf{b}_j$ , representing the  $j$ th pair of items subjects study in a list. In a convolution-correlation model (e.g., Metcalfe, 1985; Murdock, 1982, 1997), the storage equation

calls for adding the convolution of  $\mathbf{a}_j$  and  $\mathbf{b}_j$  to the memory vector,  $\mathbf{m}$ . That is,

$$\mathbf{m}_j = \mathbf{m}_{j-1} + \mathbf{a}_j * \mathbf{b}_j, \quad (4)$$

where  $j$  is the index of the current to-be-learned pair and the asterisk denotes the convolution operation.<sup>3</sup> For a list of  $L$  pairs,

$$\mathbf{m} = \sum_{j=1}^L \mathbf{a}_j * \mathbf{b}_j.$$

Associative retrieval is accomplished by correlation (denoted by #). Correlating the memory vector by a cue item,  $\mathbf{a}_k$ , retrieves a linear superposition of all items that are associated with the cue, along with error terms that represent crosstalk among items stored together in memory.<sup>4</sup> Formally, one may write

$$\mathbf{a}_k \# \mathbf{m} = \mathbf{a}_k \# (\mathbf{a}_k * \mathbf{b}_k) + \sum_{l \neq k} \mathbf{a}_k \# (\mathbf{a}_l * \mathbf{b}_l) = \mathbf{b} + \text{error}.$$

Unlike matrix models, convolution is commutative (i.e.,  $\mathbf{a} * \mathbf{b} = \mathbf{b} * \mathbf{a}$ ). This allows for no distinction between forward and backward associations. In this sense, convolution-based models implement a pure form of the ASH (Murdock, 1985; Pike, 1984).

**The search of associative memory (SAM) model.** In SAM, rehearsal of items in a limited-capacity short-term store determines the strengths of item-to-item and item-to-context associations. The  $\mathbf{A} \rightarrow \mathbf{B}$  and  $\mathbf{B} \rightarrow \mathbf{A}$  associative weights are stored separately and can be set symmetrically (e.g., Raaijmakers & Shiffrin, 1980) or biased for better forward than backward recall (e.g., Gillund & Shiffrin, 1984; Howard & Kahana, 1999; Kahana, 1996).

### Model Mimicry: Asymmetric Retrieval in Symmetric Associative Models

The large body of evidence reviewed in Part 1 suggests that the probability of successfully recalling a paired associate is largely independent of the order in which that pair was encoded. This symmetric property of associative recall is not diagnostic among the models reviewed here. This is because a model with separate forward and backward associations can mimic associative symmetry by assuming that forward and backward associations are equally weighted, on average. On the other hand, the occasional reports of asymmetric associative retrieval appear to be fatal for models assuming a symmetric associative mechanism (e.g., convolution-correlation models and autoassociative matrix models). This is because symmetric models make no distinction between forward and backward associations and are, therefore, unable to separately modify their strengths.

In this subsection, we will explore ways in which symmetric models could actually account for asymmetric retrieval, either with respect to order of encoding or with respect to differences in the properties of the  $\mathbf{A}$  and the  $\mathbf{B}$  items.

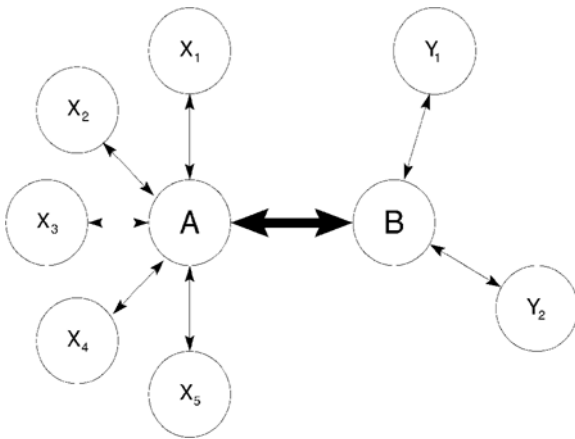


Figure 1. A simple associative network. The thick arrow depicts the experimental-episodic association between A and B. In this example, A has five preexperimental associates (marked  $X_1 \dots X_5$ ), whereas B has only two preexperimental associates (marked  $Y_1$  and  $Y_2$ ).

Consider a pair of items, A–B, studied in temporal succession. It is well known that the meaning of one word can bias the encoding of the next. This is especially true if the second word is a homonym (e.g., *bank*). If the meaning of A biases the coding of B, the association that is learned is between A and  $\tilde{B}$ , where  $\tilde{B}$  is the biased version of B. During a later test phase, the experimenter may probe for either forward or backward recall. If forward recall is probed for, the cue item, A, will readily retrieve its associate,  $\tilde{B}$ . However, cuing with B will not readily retrieve A, because the subject stored the association A– $\tilde{B}$ , rather than A–B. Such a mechanism could explain the forward-recall advantage seen in some studies (e.g., Wollen et al., 1970).

There is another mechanism that can produce asymmetric recall in models that use symmetric associative mechanisms. We show how this mechanism could explain retrieval asymmetries that arise when the to-be-associated items are drawn from different stimulus classes.

Suppose that a subject learns an A–B pair and that each item has a given number of preexperimental associates. Figure 1 illustrates the situation in which A has five preexperimental associates ( $X_1$ – $X_5$ ) and B has two preexperimental associates ( $Y_1$  and  $Y_2$ ). Each associate is unique to A or unique to B, and does not provide an indirect link between them. According to a simple associative model, cuing with either A or B will retrieve both the target item and the preexperimental associates of the cue item. The activation of these competing associations will impair retrieval. In this example, cuing with A to recall B should produce greater interference than will cuing with B to recall A. This is because A will activate five preexperimental associates, whereas B will activate only two.

One can easily formalize these ideas within the framework of a symmetric distributed memory model. For example, in an autoassociative matrix model with equally

weighted preexperimental and experimental associations, the storage equation would be given by the following:

$$W = \underbrace{(\mathbf{a} + \mathbf{b})(\mathbf{a} + \mathbf{b})'}_{\text{A} \leftrightarrow \text{B association}} + \underbrace{\sum_{i=1}^J (\mathbf{a} + \mathbf{x}_i)(\mathbf{a} + \mathbf{x}_i)'}_{\text{preexperimental associates of A}} + \underbrace{\sum_{i=1}^K (\mathbf{y}_i + \mathbf{b})(\mathbf{y}_i + \mathbf{b})'}_{\text{preexperimental associates of B}}$$

where  $W$  is the memory matrix,  $\mathbf{a}$  and  $\mathbf{b}$  are vectors representing the to-be-associated items, and  $\mathbf{x}_i$  and  $\mathbf{y}_i$  are preexperimental associates of the  $\mathbf{a}$  and  $\mathbf{b}$  items, respectively. Cuing memory with the item  $\mathbf{a}$  retrieves a linear combination of  $\mathbf{a}$ ,  $\mathbf{b}$ , and the  $J$  preexperimental associates of  $\mathbf{a}$ :

$$W\mathbf{a} = (J + 1)\mathbf{a} + \mathbf{b} + \sum_{i=1}^J \mathbf{x}_i.$$

Similarly, cuing memory with  $\mathbf{b}$  retrieves a linear combination of  $\mathbf{a}$ ,  $\mathbf{b}$ , and the  $K$  preexperimental associates of  $\mathbf{b}$ :

$$W\mathbf{b} = (K + 1)\mathbf{b} + \mathbf{a} + \sum_{i=1}^K \mathbf{y}_i.$$

Because the preexperimental associates interfere with retrieval of the target item, asymmetries will readily emerge: If  $J < K$ , retrieval from  $\mathbf{a}$  to  $\mathbf{b}$  is easier; if  $K < J$ , retrieval from  $\mathbf{b}$  to  $\mathbf{a}$  is easier; if  $J = K$ , retrieval is symmetrical (i.e.,  $\mathbf{b}$  retrieves  $\mathbf{a}$  as well as  $\mathbf{a}$  retrieves  $\mathbf{b}$ ).

This analysis applies equally well to convolution-based associative models (e.g., Metcalfe, 1985; Murdock, 1982). For example, in Metcalfe's (1985) composite holographic associative recall model (CHARM), one would write the storage equation as

$$\mathbf{m} = \underbrace{(\mathbf{a} + \mathbf{b}) * (\mathbf{a} + \mathbf{b})}_{\text{A} \leftrightarrow \text{B association}} + \underbrace{\sum_{i=1}^J (\mathbf{a} + \mathbf{x}_i) * (\mathbf{a} + \mathbf{x}_i)}_{\text{preexperimental associates of A}} + \underbrace{\sum_{i=1}^K (\mathbf{y}_i + \mathbf{b}) * (\mathbf{y}_i + \mathbf{b})}_{\text{preexperimental associates of B}}$$

where  $\mathbf{m}$  is the memory vector. Cuing memory with  $\mathbf{a}$  retrieves

$$(J + 1)\mathbf{a} + 2\mathbf{b} + 2\sum_{i=1}^J \mathbf{x}_i,$$

and cuing memory with  $\mathbf{b}$  retrieves

$$2\mathbf{a} + (K + 1)\mathbf{b} + 2\sum_{i=1}^K \mathbf{y}_i.$$

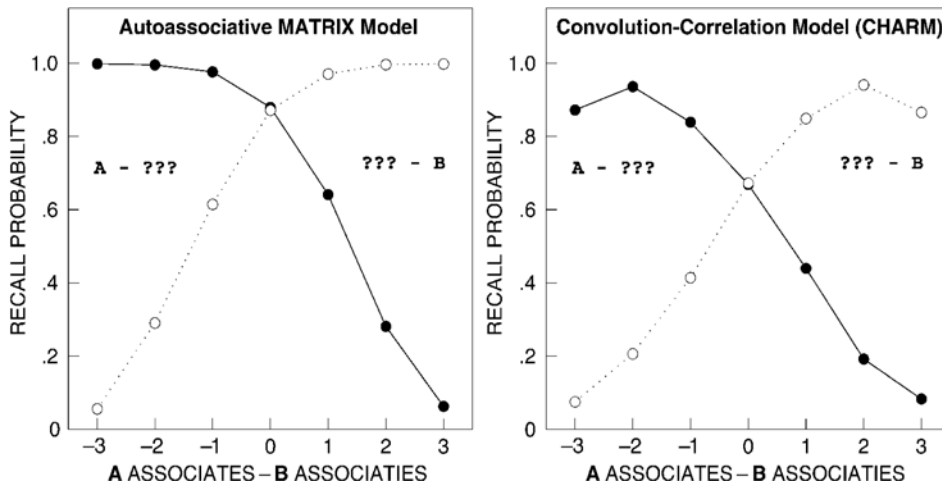


Figure 2. Model mimicry. The left panel illustrates, for an autoassociative matrix model, the effect of pre-experimental associates on retrieval. The right panel illustrates the same effects for a convolution-correlation model. Forward recall (filled circles) equals backward recall (open circles) when A and B possess the same number of preexperimental associates. When A has a greater number of preexperimental associates, retrieval of B is impaired. Conversely, when B has a greater number of preexperimental associates, recall of A is impaired.

As in the matrix model, preexperimental associates interfere with retrieval of the target item, resulting in asymmetries, if  $J \neq K$ .

Figure 2 plots the probability of forward and backward recall for a pair (A–B) whose constituent items have different numbers of preexperimental associates. The left panel shows simulation results for the autoassociative matrix model described above. The right panel shows simulation results for the CHARM model (Appendix A provides technical details of these simulations). As can be seen, when A and B possess the same number of preexperimental associates, retrieval is symmetric. However, if A has a greater number of preexperimental associates than does B, recall of B will suffer. As is shown by these simulations, associative memory models predict massive interference from preexperimental associates. Contextual information, not included in these simulations, may help to overcome these interference effects. For example, in a number of recent distributed models, items are either bound or associated with a time-varying contextual code (Howard & Kahana, 2002; Murdock, 1997). Such a temporal context may be used to selectively cue recall of information learned in temporal proximity.

Although difficulty of cued recall rarely depends on order of study, it is quite often the case that one member of a pair is a much better cue than the other member (independent of order of study). The preceding analysis shows that even symmetric models can account for pair-specific asymmetries. This suggests that it may not be necessary to postulate separate forward- and backward-recall mechanisms to explain paired-associate learning.

In the third and final part of this article, the symmetry question will be examined at the level of individual studied pairs. Using the method of successive tests, we will es-

timate the correlation between the information supporting forward and backward recall. This will allow us to directly test the independence-dependence assumptions that differentiate the ASH and the IAH.

### PART 3 Measuring the Correlation Between Forward and Backward Recall

A strict interpretation of the ASH implies that studying the pair A–B produces identical increments in A→B and B→A associative strength. Recall probabilities cannot be used to test this property. Instead, one needs to know something about the correlation between A→B and B→A associative strengths for a given studied pair. Such information can be derived only from contingency analyses applied to successive memory tests.

#### Yule's $Q$ and the Method of Successive Tests

In the successive-testing paradigm (e.g., Martin, 1971; Tulving & Thompson, 1973), subjects take two memory tests for the same or related information. In a popular application of this method (Nilsson & Gardiner, 1993; Tulving & Wiseman, 1975), subjects first study a list of A–B pairs. The experimenter first tests recognition memory for the B items, asking the subjects to distinguish these list items from lure items that had not been previously presented. On a second test, the experimenter presents each A item as a cue to recall its pair. By testing memory for B twice, first with a recognition task and then later with a recall task, one can assess the correlation between successful recognition and successful recall of a given item.

This correlation is different from a correlation across subjects or across items. In the case of a correlation across

subjects, one asks whether individuals who are good at recalling items are also good at recognizing them. In the case of a correlation across items, one asks whether items that are easily recalled are also easily recognized. Mandler (1959) demonstrated that correlating across subjects can often yield very different results than correlating across items.

Unlike correlations across subjects or items, the correlation between the outcomes of successive tests informs one of the relationship between the storage of information supporting successful recall and that supporting successful recognition (see Kahana, 2000, for a review). However, because one cannot average over subjects or items, one must compute a contingency table from the pairs of outcomes on Test 1 and Test 2 for each subject–item.

Continuing with our example, consider the correlation between recall and recognition of an item. According to a simple version of strength theory, recognition and recall tap exactly the same information, but recall requires a higher threshold. This implies that an item that is recalled will always be recognized, making the correlation between recognition and recall, at the level of subject–items, exactly 1.0. Contrary to this prediction, data from Tulving and Thompson (1973) demonstrated that subjects often fail to recognize items that they subsequently recall.

Because the test outcomes are binary variables, one computes the subject–item correlation between recognition and recall by tabulating the data in a contingency table. Yule's  $Q$ , a popular measure of correlation for  $2 \times 2$  contingency tables (Bishop, Fienberg, & Holland, 1975), is defined as  $Q = (ad - bc)/(ad + bc)$ , where  $a$  is the number of times responses on both tests were correct,  $d$  is the number of times responses on both tests were incorrect,  $b$  is the number of times responses on Test 1 were correct but those on Test 2 were not, and  $c$  is the number of times responses on Test 2 were correct but those on Test 1 were not. For a  $2 \times 2$  contingency table,  $Q$  is the same as Goodman and Kruskal's gamma statistic. Like a standard Pearson correlation,  $Q$  varies from  $-1.0$  (perfect negative correlation) to  $+1.0$  (perfect positive correlation). Together with  $Q$ , the percentage correct for Tests 1 and 2 fully characterize the data in a  $2 \times 2$  contingency table.

In the case of successive item recognition and cued-recall tests,  $Q$  is typically in the range of .4–.7, despite variations in factors that dramatically influence the probability of either recognition or recall (Nilsson & Gardiner, 1993; Tulving & Wiseman, 1975). This makes the successive-testing data inconsistent with a simple strength-threshold theory (Tulving, 1983) and also with certain distributed memory models (Kahana, Rizzuto, & Schneider, 2002).

The foregoing example illustrates the comparison of tasks, using contingencies of outcomes on successive tests. In comparing performance across the two tests, one must recognize that the tests are not measuring the same information. The first measure is of the consequences of study—subject to all the input–output interference effects

that may operate on retrieval of individual list items (e.g., Tulving & Arbuckle, 1966). The second measure is affected not only by study and the interpolated conditions, but also by the earlier test and its outcome.

One can also analyze contingencies across successive trials of the same task. This approach was fruitfully used by Estes (1960) in his studies of one-trial learning and by Tulving (1964) in his analysis of inter- versus intratrial forgetting in multitrial free recall. In this last part of the article, we will apply this technique to measure the correlation between forward and backward associative strengths in successive cued-recall tests.

### Modeling Successive Recall Tests

Consider a general heteroassociative matrix model with separate weights on forward and backward associations (see Equation 3). If the forward and backward associative strengths are perfectly correlated, this model obeys the ASH. If, however, the strengths of forward and backward associations are independent, this model exemplifies the IAH. Within this framework, the correlation between forward and backward associative learning defines a continuum spanning the ASH and the IAH.

We begin by modeling the weights on the forward and backward associations,  $\gamma_f$  and  $\gamma_b$ , respectively, as correlated random variables (with correlation given by  $\rho$ ; see Equation 3). We further assume that these variables have means of 1.0 and equal variance. Because retrieval is proportional to the match of the retrieved information with the desired target item, the probability of recalling  $\mathbf{b}_k$  is proportional to  $W\mathbf{a}_k \cdot \mathbf{b}_k$ . Similarly, the probability of recalling  $\mathbf{a}_k$  is proportional to  $W\mathbf{b}_k \cdot \mathbf{a}_k$ .

Although there is no analytic solution for the correlation between successful forward and backward recall, we can derive the correlation between the quality of information retrieved on successive recall tests. By using the variance and covariance terms derived in Appendix B and making the simplifying assumption that  $\gamma_f$  and  $\gamma_b$  have equal variance, the correlation between information driving forward and backward recall is given by the following:

$$r = \frac{\rho\sigma^2 + a_1N^{-1} + a_2N^{-2} + a_3N^{-3}}{\sigma^2 + b_1N^{-1} + b_2N^{-2} + b_3N^{-3}}, \quad (5)$$

where

$$\begin{aligned} a_1 &= 4(\rho\sigma^2 + 1), \\ a_2 &= (2L + 7)\rho\sigma^2 + 12\sigma^2 + (2L + 14), \\ a_3 &= 6\rho\sigma^2 + (4L + 2)\sigma^2 + (4L + 10), \\ b_1 &= 4(\sigma^2 + 1), \\ b_2 &= (2L + 7)\sigma^2 + 12\rho\sigma^2 + (2L + 14), \end{aligned}$$

and

$$b_3 = 6\sigma^2 + (4L + 2)\rho\sigma^2 + (4L + 10).$$

In these equations,  $L$  is the number of studied pairs,  $\sigma^2$  is the variance of  $\gamma_f$  and  $\gamma_b$ , and  $N$  is the dimensionality of



the vectors. In applications of distributed memory models to list-learning data, the value of  $N$  must be large (typically, by at least a factor of 10) in relation to  $L$ .

It may be seen that the value of  $r$  approaches that of  $\rho$  (the correlation between  $\gamma_f$  and  $\gamma_b$ ) so long as  $\sigma > 0$ .<sup>5</sup> Thus, if  $\rho = 0$  (as postulated by the IAH), the correlation between successive forward and backward recall of a given pair would be 0, whereas if  $\rho = 1$  (as postulated by the ASH), the correlation between successive forward and backward recall would be 1. In both cases, the predicted correlation between successive tests in the same direction (i.e., forward on both tests or backward on both tests) would be exactly 1, because  $\text{cov}[\mathbf{W}\mathbf{a}_k \cdot \mathbf{a}_k, \mathbf{W}\mathbf{a}_k \cdot \mathbf{a}_k] = \text{var}[\mathbf{W}\mathbf{a}_k \cdot \mathbf{a}_k]$ .

One can measure the correlation between forward and backward recall by first having subjects study a list of paired associates and later testing each pair twice, first in the forward order and then in the backward order (or first in the backward order and then in the forward order). The ASH would predict that the correlation between forward and backward recall would be 1, whereas the IAH would predict that the correlation would be 0.

### Interpretive Complexities

As with any correlational analysis, other sources of variability may influence forward and backward recall, in concert or separately. These factors may either increase or decrease the observed correlation, muddling the mapping between theory and data presented above. Consider, for example, variability in overall attention. If subjects pay attention to half of the items, ignoring the others (to take an extreme example), the attended pairs will be better recalled, and the unattended pairs worse recalled, regardless of cuing direction. This will increase the observed correlation between the successive tests. Subjects' attention during the two recall tests can also vary. Suppose that a given **A**–**B** pair was well learned. By a symmetry account, forward recall and backward recall should both succeed, because the **A**→**B** and the **B**→**A** associations are equivalent. But if a subject's attention wanders during the test, an otherwise recallable item may not be recalled, thereby lowering the expected correlation between forward and backward recall.

Furthermore, it is clear that the second test is not a direct measure of the study event. Rather, the second test reflects both the study event and the effect of the first test on memory. If successful recall on Test 1 produces further associative learning (as one might expect), recall on Test 2 will be enhanced. Humphreys and Bowyer (1980) presented a similar line of reasoning with respect to the analysis of successive recognition and recall tests (cf. Begg, 1979; Donnelly, 1988). Ideally, one would like to measure the difference in learning for remembered and nonremembered items. However, because the outcome of Test 1 is dependent on the efficiency of original learning, it is impossible to determine the effect of Test 1 encoding on the Test 1–Test 2 correlation.

Although the correlation between the outcomes of successive cued-recall tests can provide important evidence

pertaining to the question of associative symmetry, the presence of potential confounds requires one to proceed with caution (Hintzman, 1987; Kahana, 2000). Rather than focus on the level of correlation between forward and backward recall, one can compare the correlation between forward and backward recall with the correlation between forward and forward recall (or backward and backward recall). Allowing for some movement, upward or downward, in the overall level of the correlations, the ASH still predicts that the correlation between recall of pairs successively tested in incongruent (reverse) directions would be nearly identical. In contrast, the IAH predicts a higher correlation between congruent tests than between incongruent tests. This suggests that rather than focusing on the absolute level of the correlations, one is well served to compare correlations among critical conditions.

The following experiment tested these predictions. Subjects studied a list of paired associates and were then given two phases of cued-recall tests: Some pairs of words were tested in the same direction (congruent condition) in both phases, whereas other pairs were tested in reverse order (incongruent condition) in the two phases. We actually compared all four combinations of forward and backward probes on Test 1 and forward and backward probes on Test 2.

### Method

**Subjects.** Fifteen Brandeis undergraduates each participated for a payment of \$7.

**Procedure.** Figure 3 shows a schematic of the experimental design for a single experimental block. Subjects studied 12 noun pairs randomly sampled, without replacement, from the Toronto Word Pool (Friendly, Franklin, Hoffman, & Rubin, 1982). A computer displayed each pair for 2 sec, and the subjects were instructed to read each pair aloud from left to right. The number of presentations of each study pair within a given list was varied so that an equal number of pairs (four) appeared one, three, and five times. The order of the pairs in each study list was random, subject to the constraint that successively presented pairs were always unique.

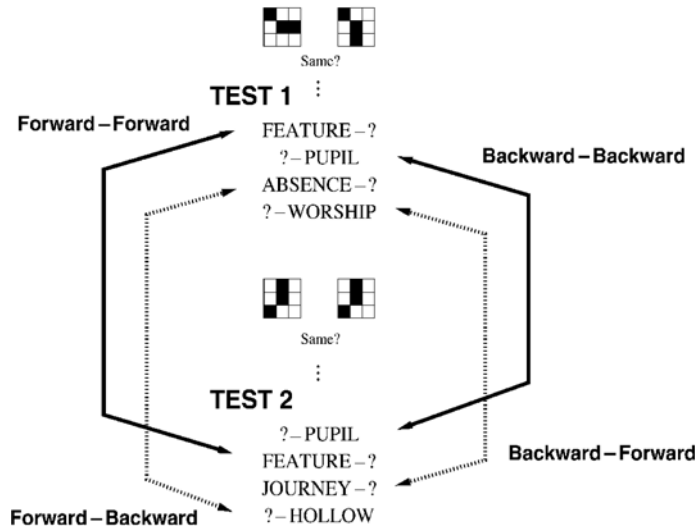
Following this study phase, the subjects performed a pattern-matching distractor task. This was done to eliminate the possibility that a given pair would be tested immediately after it was seen in the study list. The distractor task consisted of multiple trials of pattern matching. On each trial, the subjects judged whether two  $4 \times 4$  matrices were the same or different (see Figure 3). This pattern-matching task continued until the subjects correctly identified 30 consecutive pairs of matrices.

All the studied pairs were tested twice, once in each of two test phases, designated as Test 1 and Test 2. On Test 1, half of the studied pairs were cued in the forward order (**A**–?), and the other half were cued in the backward order (?–**B**). After Test 2, the subjects performed the same pattern-matching task as they did after the study list.

On Test 2, the subjects were again probed for recall of each studied pair. Among pairs tested in the forward order on Test 1, half of the pairs were tested in the forward order again on Test 2; the other half were tested in the backward order. There were thus four test conditions representing the combinations of forward and backward cues in Test 1 and Test 2: forward–forward, forward–backward, backward–forward, and backward–backward. Pairs in each repetition condition (one, three, or five presentations) were assigned to one of these four conditions. We use the term *congruent* to refer to the two same-direction conditions (forward–forward and backward–backward) and the term *incongruent* to refer to the two different-direction conditions (forward–backward and backward–forward).

**STUDY LIST**

- ABSENCE–HOLLOW
- DESPAIR–PUPIL
- JOURNEY–WORSHIP
- ABSENCE–HOLLOW
- FEATURE–COPY
- DESPAIR–PUPIL
- ABSENCE–HOLLOW
- PARTNER–MODEST
- DESPAIR–PUPIL
- ABSENCE–HOLLOW
- COMMAND–SLENDER
- SACRED–HARNES
- ABSENCE–HOLLOW



**Figure 3.** Schematic of the experimental design, illustrating the four experimental conditions (forward–forward, backward–backward, forward–backward, and backward–forward). After studying a list of pairs (some of which are shown at the top of the figure), subjects were given two successive recall tests (Test 1 and Test 2). A pattern-matching distractor task separated study, Test 1, and Test 2 phases.

On the two recall tests, each of the pairs from the study list was tested individually and in a random order. When a cue word was shown on the computer screen, the subjects attempted to vocally recall its pair from the list. Vocal responses were digitally recorded for later scoring. If subjects could not recall a target item they were instructed to say “pass.”

At the end of each block (study–Test 1–Test 2), subjects once again performed the pattern-matching task before going on to the next block.

The entire word pool was randomized separately for each subject to ensure complete randomization of all of the materials across experimental conditions. Over the course of a 1-h session, the subjects

**Table 1**  
Means and Standard Deviations of Recall Accuracy as a Function of Presentation Condition for the First and Second Successive Recall Tests

Number of Presentations	Recall Test 1				Recall Test 2			
	Forward (F–F and F–B)		Backward (B–F and B–B)		Forward (F–F and B–F)		Backward (F–B and B–B)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
1	.329	.225	.336	.222	.382	.249	.366	.213
3	.648	.291	.644	.237	.662	.281	.657	.282
5	.732	.230	.718	.209	.774	.198	.733	.229

Note—F–F designates forward recall on both Test 1 and Test 2. B–B designates backward recall on both Test 1 and Test 2. F–B designates forward recall on Test 1 and backward recall on Test 2. B–F designates backward recall on Test 1 and forward recall on Test 2.

performed six blocks consisting of list study followed by the two testing phases. No words were repeated across the different blocks.

## Results

Table 1 reports the means and standard deviations for Test 1 and Test 2 recall accuracy. Analyses of Test 1 revealed that recall improved with repetitions but did not differ between forward- and backward-recall tests. This was confirmed by a 3 (repetition, one, three, or five presentations)  $\times$  2 (recall direction, congruent vs. incongruent) repeated measures analysis of variance (ANOVA), which revealed a significant main effect for number of presentations [ $F(2,28) = 40.1, MS_e = 0.064, p < .001$ ], but not for recall direction ( $F < 1$ ). There was no significant interaction between these two factors ( $F < 1$ ). Average recall probability was .569 and .565 for forward and backward recall, respectively. This null result is consistent with the earlier literature in finding no significant difference between forward and backward recall.

Each pair that was probed on Test 2 (in either the forward or the backward direction) could have been probed in either direction on Test 1, thus producing a  $2 \times 2$  factorial of Test 2 conditions. An ANOVA on recall probability yielded the expected main effect for number of presentations [ $F(2,28) = 49.7, MS_e = 0.47, p < .001$ ]. There was also a significant crossover interaction between forward and backward probes from Test 1 to Test 2 [ $F(1,14) = 9.63, MS_e = 0.012, p < .01$ ]. As we expected on the basis of output encoding in recall (e.g., Humphreys & Bowyer, 1980; Raaijmakers & Shiffrin, 1980), the subjects performed better on Test 2 when they were tested in the same manner as they were for Test 1. This crossover interaction implies that the subjects encoded their Test 1 responses, boosting performance on Test 2. This is consistent with numerous published findings illustrating a priming or output-encoding effect of a first test on a second test. None of the other main effects or interactions approached statistical significance ( $F < 1$ ).

According to the IAH, successful recall of a given pair should be far less consistent across successive tests when cued in a different order (incongruent condition) than when cued in the same order (congruent condition). In contrast, the ASH predicts no difference in the degree of consistency across these conditions. To test these competing hypotheses, we asked whether the correlation between Test 1 and Test 2 outcomes, as measured by Yule's  $Q$ , was significantly different between congruent and incongruent conditions. In making this comparison, we separately computed  $Q$  for each subject. The mean value of  $Q$  was .88 ( $SE = .01$ ) in the congruent condition and .91 ( $SE = .02$ ) in the incongruent condition. Collapsing contingency tables across subjects, we computed  $Q$  values for each condition separately. For the combined contingency tables,  $Q$  was .99 in the congruent conditions and .97 in the incongruent conditions.<sup>6</sup> Table 2 gives the individual contingency tables for the three repetition conditions crossed with the congruent and the incongruent conditions.

As was discussed previously, it is possible that the level of correlation between successive tests is increased by variability in goodness of encoding (subjects study some pairs more than others because of overall fluctuations in attention or pair difficulty) and output encoding at Test 1. Our manipulation of number of repetitions may be useful in assessing the degree of these potential confounds. Although the three-way interaction between number of repetitions and the directionality of Tests 1 and 2 was not statistically significant, it is still conceivable that potential factors that would increase the correlation between successive tests would have their greatest effect when pairs were presented only once (as compared with three or five presentations). For the case of one presentation, the mean  $Q$  across subjects was .94 and .96 for congruent and incongruent successive tests, respectively. These values are slightly higher than those calculated when collapsing across number of repetitions. Nonetheless, the correlation does not seem to depend on the congruence between order of Test 1 and order of Test 2.

**Table 2**  
Contingency Tables, Averaged Across Subjects, for Congruent (Forward-Forward, Backward-Backward) and for Incongruent (Forward-Backward, Backward-Forward) Successive Tests

Number of Presentations		Congruent		Incongruent	
		Test 1		Test 1	
One	Test 2	+	-	+	-
				.319	.012
Three	Test 2	+	-	+	-
				.583	.012
Five	Test 2	+	-	+	-
				.729	.018
		.006	.663	.049	.537
		.037	.368	.043	.273
		.012	.241	.037	.196

Correlations as high as those reported here are rarely seen in other applications of the method of successive tests (Kahana, 2000). For example, the correlation between successive recall tests for **B** and **C** items after learning **A-B** and **A-C** associations is close to 0 (e.g., Greeno, James, & DaPolito, 1971; Martin, 1971; Martin & Greeno, 1972). For item recognition and cued-recall tests, the correlation between successive tests of the same item is typically around .55 (e.g., Kahana, 2000; Nilsson & Gardiner, 1993; Tulving & Wiseman, 1975). Our finding that the correlation between successive forward and backward recall of a given pair was nearly 1 and was no different from the correlation between successive recall in the same direction is consistent with models that assume either symmetrical or highly correlated associations, but it is not consistent with models that assume independent associations.

## GENERAL DISCUSSION

This article has reexamined a basic question about the nature of episodically formed associations in human memory. According to the IAH, separate forward and backward associations link the representations of different items in memory. Studying two items in temporal succession strengthens the forward association more than the backward association (e.g., Ebbinghaus, 1885/1913; Robinson, 1932). A strong version of this view assumes that learning produces independent increments in forward and backward associative strengths (e.g., Wolford, 1971).

According to the ASH, episodically formed verbal associations reflect the formation of a holistic conjunction of the two item representations (Asch & Ebenholtz, 1962; Murdock, 1966). Stated another way, forward and backward associations are perfectly correlated, reflecting a single underlying construct.

As reviewed in Part 1 of this article, many early studies examined retrieval asymmetries in cued recall. The goal of these studies was to establish the presence (or absence) of a forward-recall advantage. Surprisingly, investigators found nearly identical forward and backward recall under a wide range of experimental conditions. This held even when great effort was taken to ensure that subjects studied items in temporal succession. When asymmetries were found, they were typically of small magnitude (<10%) and often were not a function of order of study but, rather, of some property of the items themselves. Because the distinction between forward and backward associations was, in large part, based on the observation of these small asymmetries, some researchers claimed that the equivalence of forward and backward recall was consistent with the ASH (e.g., Asch & Ebenholtz, 1962).

These findings, however, do not allow us to select between the competing positions (e.g., Wolford, 1971). Finding that forward- and backward-recall probabilities are equal, on average, is fully consistent with either the ASH or the IAH. On the other hand, violations of symmetry

have been used to argue against the ASH (e.g., Pike, 1984). Nonetheless, neither position offers a principled explanation for the appearance of asymmetric retrieval under certain conditions.

## Mathematical Memory Models and the Problem of Model Mimicry

For the reasons cited above, research on this topic ended before it reached any conclusion. The emergence in recent years of mathematical and distributed memory models (e.g., Anderson et al., 1977; D. A. Brown et al., 1995; Chappell & Humphreys, 1994; Humphreys, Bain, & Pike, 1989; Metcalfe, 1985; Murdock, 1982, 1985, 1997; Pike, 1984) has brought this old question back into the foreground. Because these models have well-defined associative mechanisms, they allow researchers to examine the question of associative symmetry versus independent associations in more precise terms. Autoassociative matrix models and both heteroassociative and autoassociative convolution models all assume a symmetric associative process. In contrast, heteroassociative matrix models and Shiffrin's SAM model (Gillund & Shiffrin, 1984; Kahana, 1996) allow for differential weighting of forward and backward associations. It would seem that symmetric models must predict equal recall in the forward and the backward order (relative to study). In contrast, models that allow for separate forward and backward associations can just as easily account for equivalent forward- and backward-recall probabilities as they can account for asymmetric retrieval.

A more detailed analysis of these mathematical memory models shows that even models with symmetrical associative mechanisms can produce asymmetric retrieval. For example, in the models analyzed here, the effectiveness of a cue item depends on the number and strength of its preexperimental associates (relative to the target item). The item with the smallest number of unique preexperimental associates will serve as the best retrieval cue. The idea that items with fewer unique semantic associates are better retrieval cues is consistent with experimental work (Cieutat, Stockwell, & Noble, 1958). Symmetric models can also produce asymmetries in retrieval as a function of order of study. If the first item in a pair biases the coding of the second item, forward recall will be better than backward recall. This could enable these models to account for some experimental findings of order-related asymmetry (e.g., Waugh, 1970; Wollen et al., 1970).

## Stochastic Relations: A Decisive Test?

Given the possibilities of model mimicry, it was not clear how we might distinguish these two fundamentally different views of associative learning. Because recall probabilities alone did not provide an answer to this question, we sought a means of measuring the correlation between the strengths of **A**→**B** and **B**→**A** associations, for a given studied pair. Using contingency analyses applied to successive cued-recall tests, we measured the correlation between forward and backward recall of each individual

studied pair. We found that the correlation between  $A \rightarrow B$  cued recall and  $B \rightarrow A$  cued recall was comparable with the correlation between congruent successive tests (a control condition). Both correlations were close to 1.0. These findings are consistent with models that assume either symmetrical or highly correlated associations, but not with models that assume independent associations.

This new evidence seems to resolve the long-standing debate in favor of the Gestalt notion of associative symmetry. As a first approximation, one may think of associations as holistic conjunctions of meaningful representations. This conception is captured by convolution-correlation models (e.g., Metcalfe, 1985; Murdock, 1982, 1997), as well as by autoassociative matrix models (Rizuto & Kahana, 2001).

### The Generalizability of Associative Symmetry

Our findings on associative symmetry pertain to a particular experimental situation: subjects studying pairs of meaningful items and knowing that they may be tested for recall in either the forward or the backward direction. It is possible that forward and backward associations reveal their asymmetries in other experimental situations. For example, it does not seem plausible that the ASH would apply to associations among larger groups of items, as might be formed during encoding of multi-item lists. Certainly, there must be an upper limit to how many items can be reencoded as a holistic unit. For example, in free recall, following recall of an item from Position  $i$ , subjects are about twice as likely to next recall an item from Position  $i + 1$  than they are to recall an item from Position  $i - 1$  (Howard & Kahana, 1999; Kahana, 1996; Kahana, Howard, et al., 2002). Triples, on the other hand, could plausibly be encoded as holistic units or in some other form. It remains to be seen whether triples are more like pairs or more like long serial lists.

### REFERENCES

- ANDERSON, J. A., SILVERSTEIN, J. W., RITZ, S. A., & JONES, R. S. (1977). Distinctive features, categorical perception, and probability learning: Some applications of a neural model. *Psychological Review*, **84**, 413-451.
- ASCH, S. E. (1968). The doctrinal tyranny of associationism: Or what is wrong with rote learning. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory* (pp. 214-228). Englewood Cliffs, NJ: Prentice-Hall.
- ASCH, S. E., & EBENHOLTZ, S. M. (1962). The principle of associative symmetry. *Proceedings of the American Philosophical Society*, **106**, 135-163.
- BARTLING, C. A., & THOMPSON, P. (1977). Encoding specificity: Retrieval asymmetry in the recognition failure paradigm. *Journal of Experimental Psychology: Human Learning & Memory*, **3**, 690-700.
- BEGG, I. (1979). Trace loss and the recognition failure of unrecalled words. *Memory & Cognition*, **7**, 113-123.
- BISHOP, Y. M. M., FIENBERG, S. E., & HOLLAND, P. W. (1975). *Discrete multivariate analysis: Theory and practice*. Cambridge, MA: MIT Press.
- BROWN, D. A., DALLOZ, P., & HULME, C. (1995). Mathematical and connectionist models of human memory: A comparison. *Memory*, **2**, 113-145.
- BROWN, T. H., & CHATTARJI, S. (1995). Hebbian synaptic plasticity. In M. A. Arbib (Ed.), *The handbook of brain theory and neural networks* (pp. 454-459). Cambridge, MA: MIT Press.
- BUHMANN, J., DIVKO, R., & SCHULTEN, K. (1989). Associative memory with high information content. *Physical Review A*, **39**, 2689-2692.
- CHAPPELL, M., & HUMPHREYS, M. S. (1994). An auto-associative neural network for sparse representations: Analysis and application to models of recognition and cued recall. *Psychological Review*, **101**, 103-128.
- CIEUTAT, V. J., STOCKWELL, F. E., & NOBLE, C. E. (1958). The interaction of ability and amount of practice with stimulus and response meaningfulness ( $m$ ,  $m'$ ) in paired-associate learning. *Journal of Experimental Psychology*, **56**, 193-202.
- DONNELLY, R. E. (1988). Priming effects in successive episodic tests. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 256-265.
- EBBINGHAUS, H. (1885/1913). *Memory: A contribution to experimental psychology* (H. A. Ruge & C. A. Bussenius, Trans.). New York: Teachers College, Columbia University. (Original work published 1885)
- EKSTRAND, B. R. (1966). Backward associations. *Psychological Bulletin*, **65**, 50-64.
- ESTES, W. K. (1960). Learning theory and the new "mental chemistry." *Psychological Review*, **67**, 207-223.
- FLEXSER, A. J. (1981). Homogenizing the  $2 \times 2$  contingency table: A method for removing dependencies due to subject and item differences. *Psychological Review*, **88**, 327-339.
- FRIENDLY, M., FRANKLIN, P. E., HOFFMAN, D., & RUBIN, D. C. (1982). The Toronto Word Pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavior Research Methods & Instrumentation*, **14**, 375-399.
- GARDINER, J. M., GREGG, V. H., & HAMPTON, J. A. (1988). Word frequency and generation effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 687-693.
- GILLUND, G., & SHIFFRIN, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, **91**, 1-67.
- GREENO, J. G., JAMES, C. T., & DAPOLITO, F. J. (1971). A cognitive interpretation of negative transfer and forgetting of paired associates. *Journal of Verbal Learning & Verbal Behavior*, **10**, 331-345.
- HERTZ, J., KROGH, A., & PALMER, R. G. (1991). *Introduction to the theory of neural computation*. Redwood City, CA: Addison-Wesley.
- HINTZMAN, D. (1987). Recognition and recall in MINERVA 2: Analysis of the "recognition-failure" paradigm. In P. Morris (Ed.), *Modelling cognition* (pp. 215-229). New York: Wiley.
- HOPFIELD, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, **84**, 8429-8433.
- HOROWITZ, L. M., BROWN, Z. M., & WEISSBLUTH, S. (1964). Availability and the direction of associations. *Journal of Experimental Psychology*, **68**, 541-549.
- HOROWITZ, L. M., NORMAN, S. A., & DAY, R. S. (1966). Availability and associative symmetry. *Psychological Review*, **73**, 1-15.
- HOWARD, M. W., & KAHANA, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 923-941.
- HOWARD, M. W., & KAHANA, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, **46**, 269-299.
- HUMPHREYS, M. S., BAIN, J. D., & PIKE, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic, and procedural tasks. *Psychological Review*, **96**, 208-233.
- HUMPHREYS, M. S., & BOWYER, P. A. (1980). Sequential testing effects and the relationship between recognition and recognition failure. *Memory & Cognition*, **8**, 271-277.
- HUMPHREYS, M. S., PIKE, R., BAIN, J. D., & TEHAN, G. (1989). Global matching: A comparison of the SAM, Minerva II, Matrix, and TODAM models. *Journal of Mathematical Psychology*, **33**, 36-67.
- JORDAN, M. I. (1986). An introduction to linear algebra in parallel distributed processing. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Exploration in the microstructure of cognition. Vol 1: Foundations* (pp. 365-422). Cambridge, MA: MIT Press.
- KAHANA, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, **24**, 103-109.
- KAHANA, M. J. (2000). Contingency analyses of memory. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of human memory* (pp. 323-384). Oxford: Oxford University Press.
- KAHANA, M. J., & CAPLAN, J. B. (2002). Associative asymmetry in probed recall of serial lists. *Memory & Cognition*, **30**, 841-849.

- KAHANA, M. J., HOWARD, M. W., ZAROMB, F., & WINGFIELD, A. (2002). Age dissociates recency and lag-recency effects in free recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 530-540.
- KAHANA, M. J., RIZZUTO, D. S., & SCHNEIDER, A. (2002). *An analysis of the recognition-recall relation in four distributed memory models*. Manuscript submitted for publication.
- KÖHLER, W. (1947). *Gestalt psychology*. New York: Liveright.
- LEVY, C. M., & NEVILL, D. D. (1974). B-A learning as a function of degrees of A-B learning. *Journal of Experimental Psychology*, **102**, 327-329.
- LOCKHART, R. S. (1969). Retrieval asymmetry in the recall of adjectives and nouns. *Journal of Experimental Psychology*, **79**, 12-17.
- MANDLER, G. (1959). Stimulus variables and subject variables: A caution. *Psychological Review*, **66**, 145-149.
- MANDLER, G., RABINOWITZ, J. C., & SIMON, R. A. (1981). Coordinate organization: The holistic representation of word pairs. *American Journal of Psychology*, **92**, 209-222.
- MARTIN, E. (1971). Verbal learning theory and independent retrieval phenomena. *Psychological Review*, **78**, 314-332.
- MARTIN, E., & GREENO, J. G. (1972). Independence of associations tested. *Psychological Review*, **79**, 265-267.
- MCNAUGHTON, B. L., & MORRIS, R. G. (1987). Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends in Neurosciences*, **10**, 408-415.
- METCALFE, J. (1985). Levels of processing, encoding specificity, elaboration, and CHARM. *Psychological Review*, **92**, 1-38.
- METCALFE, J. (1991). Recognition failure and the composite memory trace in CHARM. *Psychological Review*, **98**, 529-553.
- METCALFE, J. (1993). Novelty monitoring, metacognition, and control in a composite holographic associative recall model: Implications for Korsakoff amnesia. *Psychological Review*, **100**, 3-22.
- MURDOCK, B. B. (1956). Backward learning in paired associates. *Journal of Experimental Psychology*, **51**, 213-215.
- MURDOCK, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, **64**, 482-488.
- MURDOCK, B. B. (1965). Associative symmetry and dichotic presentation. *Journal of Verbal Learning & Verbal Behavior*, **4**, 222-226.
- MURDOCK, B. B. (1966). Forward and backward associations in paired associates. *Journal of Experimental Psychology*, **71**, 732-737.
- MURDOCK, B. B. (1979). Convolution and correlation in perception and memory. In L.-G. Nilsson (Ed.), *Perspectives in memory research: Essays in honor of Uppsala University's 500th anniversary* (pp. 105-119). Hillsdale, NJ: Erlbaum.
- MURDOCK, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, **89**, 609-626.
- MURDOCK, B. B. (1985). Convolution and matrix systems: A reply to Pike. *Psychological Review*, **92**, 130-132.
- MURDOCK, B. B. (1989). Learning in a distributed memory model. In C. Izawa (Ed.), *Current issues in cognitive processes: The Floweree Symposium on Cognition* (pp. 69-106). Hillsdale, NJ: Erlbaum.
- MURDOCK, B. B. (1995). Similarity in a distributed memory model. *Journal of Mathematical Psychology*, **39**, 251-264.
- MURDOCK, B. B. (1997). Context and mediators in a theory of distributed associative memory (TODAM2). *Psychological Review*, **104**, 839-862.
- NAIRNE, J. S., & WIDNER, R. L. (1988). Familiarity and lexicality as determinants of the generation effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 694-699.
- NILSSON, L.-G., & GARDINER, J. M. (1993). Identifying exceptions in a database of recognition failure studies from 1973 to 1992. *Memory & Cognition*, **21**, 397-410.
- PIKE, R. (1984). Comparison of convolution and matrix distributed memory systems for associative recall and recognition. *Psychological Review*, **91**, 281-294.
- PRIMOFF, E. (1938). Backward and forward associations as an organizing act in serial and in paired-associate learning. *Journal of Psychology*, **5**, 375-395.
- RAAIJMAKERS, J. G. W., & SHIFFRIN, R. M. (1980). SAM: A theory of probabilistic search of associative memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 14, pp. 207-262). New York: Academic Press.
- RIZZUTO, D. S., & KAHANA, M. J. (2001). An autoassociative model of paired associate learning. *Neural Computation*, **13**, 2075-2092.
- ROBINSON, E. S. (1932). *Association theory to-day; an essay in systematic psychology*. New York: Century.
- ROCK, I. (1957). The role of repetition in associative learning. *American Journal of Psychology*, **70**, 186-193.
- ROCK, I., & CERASO, J. (1964). Toward a cognitive theory of associative learning. In C. Scheerer (Ed.), *Cognition: Theory, research and promise* (pp. 110-146). New York: Harper & Row.
- SLAMECKA, N. J. (1976). An analysis of double-function lists. *Memory & Cognition*, **4**, 581-585.
- TREVES, A., & ROLLS, E. T. (1994). Computational analysis of the role of the hippocampus in memory. *Hippocampus*, **4**, 374-391.
- TULVING, E. (1964). Intratrial and intertrial retention: Notes towards a theory of free recall verbal learning. *Psychological Review*, **71**, 219-237.
- TULVING, E. (1983). *Elements of episodic memory*. New York: Oxford University Press.
- TULVING, E., & ARBUCKLE, T. Y. (1966). Input and output interference in short-term associative memory. *Journal of Experimental Psychology*, **72**, 145-150.
- TULVING, E., & THOMPSON, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, **80**, 352-373.
- TULVING, E., & WISEMAN, S. (1975). Relation between recognition and recognition failure of recallable words. *Bulletin of the Psychonomic Society*, **6**, 79-82.
- WARD, G., & WOODWARD, G. (2002). *Three factors that cause the word frequency effect in free recall*. Manuscript submitted for publication.
- WAUGH, N. C. (1970). Associative symmetry and recall latencies: A distinction between learning and performance. *Acta Psychologica*, **33**, 326-337.
- WEBER, E. U. (1988). Expectation and variance of item resemblance distributions in a convolution-correlation model of distributed memory. *Journal of Mathematical Psychology*, **32**, 1-43.
- WOLFORD, G. (1971). Function of distinct associations for paired-associate performance. *Psychological Review*, **78**, 303-313.
- WOLLEN, K. A., FOX, R. A., & LOWRY, D. H. (1970). Variations in asymmetry as a function of degree of forward learning. *Journal of Verbal Learning & Verbal Behavior*, **86**, 416-419.

## NOTES

1. There were several problems with the studies reported by Asch and Ebenholtz (1962). First, they used a moderately long presentation rate (3 sec for A, 3 sec for A-B). At this rate, rehearsal strategies may play a significant role in the storage process. If subjects rehearse pairs in an A-B-A-B fashion, incidental learning of B→A associations may occur. Second, Asch and Ebenholtz's experiments had low statistical power: Even in their critical experiment demonstrating symmetry, their comparison was based on 96 observations (12 subjects each studied 8 pairs of nonsense syllables for forward- and backward-recall tests).

2. Introducing a nonlinear dynamical rule enables the network to settle into the correct retrieval state even when items are correlated (see Hertz, Krogh, & Palmer, 1991, for a review).

3. If each of two item vectors has  $N$  elements, the convolution will be a vector with  $2N - 1$  elements. The convolution of  $\mathbf{a}$  and  $\mathbf{b}$  is defined by the equation  $\mathbf{a} * \mathbf{b}(m) = \sum_i \mathbf{a}(i) \mathbf{b}(m - i)$ ; where  $m$  is the index to the elements in the convolution vector and  $i$  indexes the elements in the item vectors  $\mathbf{a}$  and  $\mathbf{b}$ .

4. The correlation of  $\mathbf{a}$  and  $\mathbf{b}$  is defined by the equation  $\mathbf{a} \# \mathbf{b}(m) = \sum_i \mathbf{a}(i) \mathbf{b}(m + i)$ , where  $m$  is the index to the elements of the correlation vector and  $i$  indexes the elements of the item vectors. One should note that correlation is like convolution, except for a change of sign in the summation.

5. Without sufficient variability in the models' learning parameters, the correlation between forward and backward recall rises above the predicted value  $\rho$ ; and, in the limit as  $\sigma$  approaches zero,  $r$  approaches 1.0.

6. This increase in the correlations is consistent with the well-known finding that  $Q$  increases after collapsing contingency tables (e.g., Flexser, 1981).

## APPENDIX A

### Details of the Simulations

We used Monte Carlo methods to simulate associative learning and cued recall of related word pairs in two distributed memory models: an autoassociative matrix model and a convolution-correlation model (CHARM). We asked how the number of “preexperimental” associates of the **A** and **B** items would influence performance in each of the two models.

Three hundred one element vectors represented list items in both models. To model interitem similarity, we constructed each item vector by summing a shared prototype vector with a uniquely chosen random vector. The relative weighting of these two vectors determined the interitem similarity (see Murdock, 1995, for details). Elements of the constituent vectors were independent and identically distributed (i.i.d.) Gaussian variables with a mean of zero and a variance of  $1/301$ . In these simulations, a lexicon of 50 items was constructed with interitem similarity of  $\rho = .4$ .

A weight matrix (for the matrix model) and a memory vector (for the convolution-correlation model) were initialized to zero at the beginning of the simulations. In both models, a pair of items is represented by the sum of the constituent item vectors. In the autoassociative convolution-correlation model, a pair is stored in memory by adding the autoconvolution of the pair to the memory vector. In the matrix model, the outer product of the pair with itself is added to the weight matrix.

A single pair of item vectors (**a**–**b**) was randomly chosen from the lexicon. In the preexperimental phase, the **a** item was paired with each of  $j$  item vectors, and the **b** item was paired with each of  $k$  item vectors. These pairs were stored in the memory structure (either vector or matrix, depending on the model). Prior to forming the association between the pairs (as was described above), each item was encoded probabilistically (see Murdock, 1989, for details). Probabilistic encoding means that each element in the vector is encoded perfectly with a probability of  $p$  and set to zero with a probability of  $1 - p$ . For the preexperimental items,  $p$  was set to 1.0, whereas for the experimental association between **a** and **b**,  $p$  was set to .9.

In the experimental phase, the association between **a** and **b** was stored. Then **a** was used to cue memory for **b**, and **b** was used to cue memory for **a**. Retrieval was successful if the retrieved trace had a higher match (as measured by the cosine of the angle between the vectors) with the target (**a** + **b**) than with any of the other traces in the lexicon. The entire process described above was repeated 2,000 times for each combination of  $j$  and  $k$  (the number of preexperimental associates to the **a** and **b** items) shown in Figure 2.

## APPENDIX B

### Derivations for the Bidirectional Matrix Model

We considered a heteroassociative matrix with correlated weights for forward and backward associative matrices. The storage equation for a list of  $L$  pairs is

$$W = \sum_{l=1}^L (\gamma_f \mathbf{b}_l \mathbf{a}_l' + \gamma_b \mathbf{a}_l \mathbf{b}_l')$$

The parameters for forward and backward recall,  $\gamma_f$  and  $\gamma_b$ , respectively, are correlated random variables, such that

$$E[\gamma_f] = E[\gamma_b] = 1, \text{var}[\gamma_f] = \sigma_{\gamma_f}^2, \text{var}[\gamma_b] = \sigma_{\gamma_b}^2$$

and

$$\text{cov}[\gamma_f, \gamma_b] = \rho \sigma_{\gamma_f} \sigma_{\gamma_b}.$$

The components of the random vectors, **a** and **b**, are i.i.d. Gaussian random variables with a mean of zero and a variance of  $1/N$ . The probabilities of forward and backward recall are proportional to  $W \mathbf{a}_k \cdot \mathbf{b}_k$  and  $W \mathbf{b}_k \cdot \mathbf{a}_k$ , respectively. Expanding these terms for the forward-recall case yields

$$\begin{aligned} W \mathbf{a}_k \cdot \mathbf{b}_k &= \sum_{l=1}^L (\gamma_f \mathbf{b}_l \mathbf{a}_l' + \gamma_b \mathbf{a}_l \mathbf{b}_l') \mathbf{a}_k \cdot \mathbf{b}_k \\ &= \sum_{l=1}^L [\gamma_f (\mathbf{a}_l \cdot \mathbf{a}_k) (\mathbf{b}_l \cdot \mathbf{b}_k) + \gamma_b (\mathbf{a}_l \cdot \mathbf{b}_k) (\mathbf{a}_k \cdot \mathbf{b}_l)] \\ &= \gamma_f \mathbf{T}_1 + \gamma_f \mathbf{T}_2 + \gamma_b \mathbf{T}_3 + \gamma_b \mathbf{T}_4, \end{aligned} \tag{B1}$$

APPENDIX B (Continued)

where

$$\begin{aligned} \mathbf{T}_1 &= \sum_{i=1}^N \sum_{j=1}^N [\mathbf{a}_k^2(i) \mathbf{b}_k^2(j)], \\ \mathbf{T}_2 &= \sum_{l \neq k}^L \sum_{i=1}^N \sum_{j=1}^N [\mathbf{a}_l(i) \mathbf{a}_k(i) \mathbf{b}_l(j) \mathbf{b}_k(j)], \\ \mathbf{T}_3 &= \sum_{i=1}^N \sum_{j=1}^N [\mathbf{a}_k(i) \mathbf{b}_k(i) \mathbf{a}_k(j) \mathbf{b}_k(j)], \end{aligned}$$

and

$$\mathbf{T}_4 = \sum_{l \neq k}^L \sum_{i=1}^N \sum_{j=1}^N [\mathbf{a}_l(i) \mathbf{b}_k(i) \mathbf{a}_k(j) \mathbf{b}_l(j)].$$

For the backward-recall case,

$$W \mathbf{b}_k \cdot \mathbf{a}_k = \gamma_b \mathbf{T}_1 + \gamma_b \mathbf{T}_2 + \gamma_f \mathbf{T}_3 + \gamma_f \mathbf{T}_4. \tag{B2}$$

VARIANCE AND COVARIANCE COMPONENTS

In the derivations that follow, we use  $W, X, Y,$  and  $Z$  to denote i.i.d. Gaussian variables with  $\mu = 0$ . For even powers of  $X$ , the expectations are given by  $E[X^2] = \sigma^2, E[X^4] = 3\sigma^4,$  and  $E[X^6] = 15\sigma^6$ ; for odd powers of  $X$ , the expectation is zero (see Weber, 1988, for details).

To derive the correlation between forward and backward recall in the model, we first derive the values of the nonzero components that enter into the calculations of the variance and covariance terms. These components are designated by letters in Tables B1 and B2, and the contributions to each of these components are derived in this next section.

Variance Components

Cases for A

$i \neq i', j \neq j'$ :

$$(N^4 - 2N^3) \text{cov}[\gamma_f X^2 Y^2, \gamma_f W^2 Z^2] = (1 - 2N^{-1}) \sigma_{\gamma_f}^2.$$

$i = i', j \neq j'$  or  $j = j', i \neq i'$ :

$$(2N^3 - N^2) \text{cov}[\gamma_f X^2 Y^2, \gamma_f X^2 Z^2] = (2N^{-1} - N^{-2})(3\sigma_{\gamma_f}^2 + 2).$$

$i = i', j = j'$ :

$$N^2 \text{cov}[\gamma_f X^2 Y^2, \gamma_f X^2 Y^2] = N^{-2}(9\sigma_{\gamma_f}^2 + 8).$$

Cases for B

$i = i', j = j', l = l'$ :

$$(L-1)N^2 \text{cov}[\gamma_f WXYZ, \gamma_f WXYZ] = (L-1)(\sigma_{\gamma_f}^2 + 1)N^{-2}.$$

Table B1

Variance Components for Terms in Equations 3 and 4

	$f\mathbf{T}_1$	$f\mathbf{T}_2$	$b\mathbf{T}_3$	$b\mathbf{T}_4$
$f\mathbf{T}_1$	A	0	E	0
$f\mathbf{T}_2$	0	B	0	F
$b\mathbf{T}_3$	E	0	C	0
$b\mathbf{T}_4$	0	F	0	D

Table B2

Covariance Components for Terms in Equations 3 and 4

	$b\mathbf{T}_1$	$b\mathbf{T}_2$	$f\mathbf{T}_3$	$f\mathbf{T}_4$
$f\mathbf{T}_1$	A	0	G	0
$f\mathbf{T}_2$	0	B	0	H
$b\mathbf{T}_3$	E	0	C	0
$b\mathbf{T}_4$	0	F	0	D



## APPENDIX B (Continued)

**Cases for C**

$$i = j = i' = j':$$

$$N \operatorname{cov}[\gamma_b X^2 Y^2, \gamma_b X^2 Y^2] = N^{-3} (9E[\gamma_b^2] - 1).$$

$$i = j \neq i' = j':$$

$$(N^2 - N) \operatorname{cov}[\gamma_b X^2 Y^2, \gamma_b W^2 Z^2] = (N^{-2} - N^{-3}) (E[\gamma_b^2] - 1).$$

$$i = i', j = j', i \neq j, i' \neq j', \text{ or } j = i', i = j', i \neq j, i' \neq j':$$

$$2(N^2 - N) \operatorname{cov}[\gamma_b WXYZ, \gamma_b WXYZ] = 2(N^{-2} - N^{-3}) (E[\gamma_b^2]).$$

**Cases for D**

$$i = i', j = j', l = l':$$

$$(L-1)N^2 \operatorname{cov}[\gamma_b WXYZ, \gamma_b WXYZ] = (L-1)(\sigma_{\gamma_b}^2 + 1)N^{-2}.$$

**Cases for E**

$$j \neq i = i' = j' \text{ or } i \neq j = i' = j':$$

$$2(N^2 - N) \operatorname{cov}[\gamma_f X^2 Y^2, \gamma_b X^2 Z^2] = 2(N^{-2} - N^{-3}) (3E[\gamma_f \gamma_b] - 1).$$

$$i = j = i' = j':$$

$$N \operatorname{cov}[\gamma_f X^2 Y^2, \gamma_b X^2 Y^2] = N^{-3} (9E[\gamma_f \gamma_b] - 1).$$

**Cases for F**

$$i = j = i' = j':$$

$$N(L-1) \operatorname{cov}[\gamma_f WXYZ, \gamma_b WXYZ] = 2N^{-3} (L-1) E[\gamma_f \gamma_b].$$

Using Equation B1, we can see that

$$\begin{aligned} \operatorname{var}[\mathbf{W}\mathbf{a}_k \cdot \mathbf{b}_k] &= \sigma_{\gamma_f}^2 + 4[\sigma_{\gamma_f}^2 + 1]N^{-1} \\ &\quad + [(L+5)\sigma_{\gamma_f}^2 + (L+2)\sigma_{\gamma_b}^2 + 12\rho\sigma_{\gamma_f}\sigma_{\gamma_b} + (2L+14)]N^{-2} \\ &\quad + [6\sigma_{\gamma_b}^2 + (4L+2)\rho\sigma_{\gamma_f}\sigma_{\gamma_b} + (4L+10)]N^{-3} \end{aligned}$$

and

$$\begin{aligned} \operatorname{var}[\mathbf{W}\mathbf{b}_k \cdot \mathbf{a}_k] &= \sigma_{\gamma_b}^2 + 4[\sigma_{\gamma_b}^2 + 1]N^{-1} \\ &\quad + [(L+5)\sigma_{\gamma_b}^2 + (L+2)\sigma_{\gamma_f}^2 + 12\rho\sigma_{\gamma_f}\sigma_{\gamma_b} + (2L+14)]N^{-2} \\ &\quad + [6\sigma_{\gamma_f}^2 + (4L+2)\rho\sigma_{\gamma_f}\sigma_{\gamma_b} + (4L+10)]N^{-3}. \end{aligned}$$

**Covariance Components****Cases for A**

$$i \neq i', j \neq j':$$

$$(N^4 - 2N^3) \operatorname{cov}[\gamma_f X^2 Y^2, \gamma_b W^2 Z^2] = (1 - 2N^{-1}) (\rho\sigma_{\gamma_f}\sigma_{\gamma_b}).$$

$$i = i', j \neq j' \text{ or } j = j', i \neq i':$$

$$(2N^3 - N^2) \operatorname{cov}[\gamma_f X^2 Y^2, \gamma_b X^2 Z^2] = (2N^{-1} - N^{-2}) (3\rho\sigma_{\gamma_f}\sigma_{\gamma_b} + 2).$$

$$i = i', j = j':$$

$$N^2 \operatorname{cov}[\gamma_f X^2 Y^2, \gamma_b X^2 Y^2] = N^{-2} (9\rho\sigma_{\gamma_f}\sigma_{\gamma_b} + 8).$$

## APPENDIX B (Continued)

**Cases for B**

$$i = i', j = j', l = l':$$

$$(L-1)N^2 \text{cov}[\gamma_f WXYZ, \gamma_b WXYZ] = (L-1)(\rho\sigma_{\gamma_f}\sigma_{\gamma_b} + 1)N^{-2}.$$

**Cases for C**

$$i = j = i' = j':$$

$$N \text{cov}[\gamma_b X^2 Y^2, \gamma_f X^2 Y^2] = N^{-3}(9\rho\sigma_{\gamma_f}\sigma_{\gamma_b} + 8).$$

$$i = j \neq i' = j':$$

$$(N^2 - N) \text{cov}[\gamma_b X^2 Y^2, \gamma_f W^2 Z^2] = (N^2 - N^{-3})(\rho\sigma_{\gamma_f}\sigma_{\gamma_b}).$$

$$i = i', j = j', i \neq j, i' \neq j', \text{ or } j = i', i = j', i \neq j, i' \neq j':$$

$$2(N^2 - N) \text{cov}[\gamma_b WXYZ, \gamma_f WXYZ] = 2(N^2 - N^{-3})(\rho\sigma_{\gamma_f}\sigma_{\gamma_b} + 1).$$

**Cases for D**

$$i = i', j = j', l = l':$$

$$(L-1)N^2 \text{cov}[\gamma_b WXYZ, \gamma_f WXYZ] = (L-1)N^{-2}(\rho\sigma_{\gamma_f}\sigma_{\gamma_b} + 1).$$

**Cases for E, G**

$$j \neq i = i' = j' \text{ or } i \neq j = i' = j':$$

$$2(N^2 - N) \text{cov}[\gamma_{f,b} X^2 Y^2, \gamma_{f,b} X^2 Z^2] = 2(N^2 - N^{-3})(3\sigma_{\gamma_{f,b}}^2 + 2).$$

$$i = j = i' = j':$$

$$N \text{cov}[\gamma_{f,b} X^2 Y^2, \gamma_{f,b} X^2 Y^2] = N^{-3}(9\sigma_{\gamma_{f,b}}^2 + 8).$$

**Cases for F, H**

$$i = j = i' = j':$$

$$N(L-1) \text{cov}[\gamma_{f,b} WXYZ, \gamma_{f,b} WXYZ] = 2(L-1)(\sigma_{\gamma_{f,b}}^2 + 1)N^{-3}.$$

The covariance of forward and backward recall is thus given by

$$\begin{aligned} \text{cov}[\mathbf{W}\mathbf{a}_k \cdot \mathbf{b}_k, \mathbf{W}\mathbf{b}_k \cdot \mathbf{a}_k] &= \rho\sigma_{\gamma_f}\sigma_{\gamma_b} + 4(\rho\sigma_{\gamma_f}\sigma_{\gamma_b} + 1)N^{-1} \\ &\quad + [(2L+7)\rho\sigma_{\gamma_f}\sigma_{\gamma_b} + 6(\sigma_{\gamma_b}^2\sigma_{\gamma_f}^2) + (2L+14)]N^{-2} \\ &\quad + [6\rho\sigma_{\gamma_f}\sigma_{\gamma_b} + (2L+1)(\sigma_{\gamma_b}^2\sigma_{\gamma_f}^2) + (4L+10)]N^{-3}. \end{aligned}$$

(Manuscript received by April 19, 1999;  
revision accepted for publication May 29, 2002.)