

Annu. Rev. Psychol. 1992. 43:205-34
Copyright © 1992 by Annual Reviews Inc. All rights reserved

MODELS FOR RECALL AND RECOGNITION

Jeroen G. W. Raaijmakers

TNO Institute for Perception, 3769 ZG Soesterberg, The Netherlands

Richard M. Shiffrin

Department of Psychology, Indiana University, Bloomington, Indiana 47405

KEY WORDS: mathematical models, recall, recognition, connectionism, associative memory

CONTENTS

INTRODUCTION	205
<i>Theoretical Approaches</i>	206
<i>Network Models</i>	206
<i>Episodic Trace Models</i>	207
<i>Distributed Memory Models</i>	209
<i>Differences and Similarities</i>	214
APPLICATIONS TO PARADIGMS	222
<i>Cued Recall</i>	222
<i>List Length</i>	222
<i>Interference and Forgetting</i>	222
<i>Free Recall</i>	223
<i>Recognition</i>	226
<i>Pair Recognition</i>	226
<i>List Length vs List Strength</i>	227
<i>Evaluating the Models</i>	229
<i>Number of Parameters</i>	229
<i>Number of Processes</i>	230
<i>Quantitative, Qualitative and Semi-Quantitative Fits</i>	231

INTRODUCTION

Research on learning and memory has been driven by models since at least the 1940s. Over the years, the emphasis in mathematical modeling has shifted

from precise fitting of single experiments to what might be best described as semi-quantitative fitting of a wide variety of phenomena from a number of experimental paradigms [compare for instance Bower's one-element model (Bower 1961) to any of the current models such as ACT* (Anderson 1983b), SAM (Raaijmakers & Shiffrin 1981; Gillund & Shiffrin 1984) or TODAM (Murdock 1982)].

The complexities of recent models contributes to the apparent impossibility of deciding between them. Although couched in quite different terms, they often make very similar predictions, at least under appropriate choices of parameters. This makes it difficult to generate critical empirical tests. On the other hand, the similarity of predictions suggests real progress in theory development, forced by the necessity to account for a standard and agreed upon corpus of findings.

In this chapter, we review a number of the most important contemporary models of memory, trying to highlight the similarities and differences in the way they handle basic facts about recall and recognition. Space limitations prevent us from any attempt at exhaustive coverage. For the same reason, although a number of models can or do predict response latencies, we leave coverage of this important topic to a future chapter.

Theoretical Approaches

Although all classifications are to some degree unsatisfactory, we group current models into three categories: 1. separate-trace models involving spreading activation, or making no explicit activation assumptions—we term these *network models*; 2. separate-trace models involving parallel activation, here termed *episodic trace models*; and 3. *composite/distributed memory models*.

Network Models

Network models propose that long-term memory consists of a set of nodes with links connecting the nodes. The nodes represent concepts or cognitive units (Anderson 1983a,b), the links semantic or episodic relations. Whenever two items are studied together, a link between the nodes representing these items may be formed. In most of these models, a process of spreading activation determines the retrieval of information from memory. Basically, there are two types of network model: (a) the *all-or-none activation* model, and (b) the *continuous activation* model.

The all-or-none activation model assumes that network nodes are either active or inactive. The best known example is the ACTE model proposed by Anderson (1976). In such a model, the spreading of activation is determined by the (relative) strength of the nodes or the links. Suppose that two nodes, X and Y, are connected by a link *l*. If node X is active, the probability of

activating Y in the next unit of time is a function of s/S , the relative strength of the link l compared to all other links emanating from X , or, alternatively, the relative strength of node Y compared to all other nodes linked to X . In such a model, the probability of retrieving Y , given that X is active, is equal to the likelihood that Y is activated before a specified cutoff time.

A continuous activation model was developed by Anderson (1983a,b) as an alternative to the all-or-none model. The basic difference is that network nodes now have a continuously varying activation strength. This means that one needs a different rule for determining whether a memory trace has been successfully retrieved. If a stimulus node X has an associative link to another node Y , some activation will spread from X to Y . The amount of activation of Y is determined by the relative strength of the link between X and Y (compared to all other links from X). In such a model it becomes more natural to assume that the probability and latency of retrieving the trace Y are a function of the amount of activation of Y . Thus, the notion of spreading activation has changed from gradually activating connected nodes (i.e. distant nodes take longer to activate) to a dynamic model in which the activation spreads rapidly over the network but in varying degrees (i.e. distant nodes have a lower level of activation).

As an example, in the most recent version of Anderson's ACT theory, the ACT* model (Anderson 1983b), it is assumed that during storage memory traces (called cognitive units) are formed. Traces vary in strength (a function of the number of presentations and the retention interval), and these strengths determine the amount of activation that converges on the trace from associated nodes (thus, in this model, it is relative node strength, not link strength, that determines the flow of activation; it is not evident whether this makes a difference). Thus, in a paired-associate recall situation, where the subject learns a list of pairs $A-B$, it is assumed that the trace (the cognitive unit) encodes the information that this pair was presented in this context. At test, the response will be retrieved if (a) such a trace has indeed been formed, and (b) it can be retrieved within the cutoff time.

Episodic Trace Models

The basic characteristic of episodic trace models is that they assume a set of separately stored memory traces that are activated in parallel. Such models are sometimes called "search models" because recall requires that some one of these traces must be "found" and output. In one subclass of models, recall of information from long-term memory involves sequential samples from a set of memory traces. The best-known example of such a model is the Search of Associative Memory (SAM) model proposed by Raaijmakers & Shiffrin (1980, 1981). In SAM the sampling probability of a particular trace depends on the relative strength of that trace compared to all other memory traces.

The SAM model assumes that during storage, information is represented in “memory images,” which contain item, associative, and contextual information. The amount and type of information stored are determined by coding processes in short-term store (STS). In most (intentional) learning paradigms the amount of information stored is a function of the length of time that the item is studied while in STS. According to the SAM model, retrieval from long-term store (LTS) is based on cues (context, items, category names). Whether an image is retrieved or not depends on the associative strengths of the retrieval cues to that image. These strengths are a function of the overlap of the cue information and the information stored in the image.

An important property of the SAM model is that it incorporates a rule to describe the overall strength of a set of probe cues to a particular image: The overall activation strength ($A(i)$) is equal to product of the individual cue strengths (weighted if necessary for relative salience or importance). This multiplicative feature focuses the search process on those images that are strongly associated to all cues.

In recall tasks, the search process of the SAM model is based on a series of elementary retrieval attempts. Each attempt involves selecting or sampling one image based on the relative activation strengths. Sampling an image allows recovery of information from it. For simple recall tasks, the probability of successfully recovering the name of the encoded word is assumed to be a simple function of the weighted strengths.

Although the SAM model assumes that the process of activating information is basically the same in recall and recognition, it postulates some important differences between these two processes. It is assumed that recognition does not necessarily involve sequential sampling but is (mostly) based on a direct-access process involving a single retrieval step (Gillund & Shiffrin 1984:55–56). The recognition decision in this direct access process is based on the sum ($\sum A(k)$) of the activation strengths; if the same cues are used to probe memory for recall and recognition, the activations are the same in both cases, though used in different ways. As we shall see, the process of summing activations makes the SAM model for recognition remarkably similar in structure to models that appear quite different on the surface, even models (e.g. most composite, distributed models) that sum inputs at storage rather than retrieval.

Because an “old” response is made when $\sum A(k)$ is greater than a criterion value, the distribution of the sum determines performance. For this reason, SAM incorporates specific variance assumptions; in particular, the standard deviation of the distribution of a given strength is assumed to be proportional to the mean strength value (Gillund & Shiffrin 1984; Shiffrin et al 1990).

The SAM model assumes that for typical episodic-memory tasks, contextual information is always encoded in the memory image, and context is

one of the retrieval cues. Mensink & Raaijmakers (1988, 1989) proposed an extension of the SAM model to handle time-dependent changes in context. The basic idea, adapted from Stimulus Sampling Theory (Estes 1955), is that a random fluctuation of elements occurs between two sets, a set of available context elements and a set of (temporarily) unavailable context elements. Performance is a function of the relationship between sets of available elements at different points in time (i.e. study and test trials).

Hintzman (1984, 1986, 1988) developed a model for episodic memory similar to Gillund & Shiffrin's SAM model for recognition. This model, MINERVA 2, has been applied primarily to category learning and recognition memory. It is assumed that each experience produces a separate memory trace. Both items and memory traces are represented as lists of features or vectors. In simulations of the model, it has been assumed that each feature is independently encoded with probability L , a learning rate parameter. When a probe cue is presented, all memory traces are activated in parallel. The amount of activation of any particular trace is a nonlinear function of the similarity to the probe cue.

As in the SAM model, recognition performance in MINERVA 2 depends on a single value, the summed activation of all traces. In order to allow recall to be carried out, the model also stipulates that a vector is retrieved. This vector (called the *echo*) is the sum of all trace vectors, each weighted by its activation value. Because of the weighting, and the nonlinear activation rule, the echo will contain a disproportional representation of those traces similar to the memory probe. Thus, if part of trace j is used as a probe, the echo will contain a strong representation from the entire trace j . For example, if a trace encodes a studied pair $A-B$, and A is used as a probe, the echo will contain something similar to $A-B$, allowing B to be recalled. Of course, the retrieved trace is actually a composite of many traces (unlike the SAM model), so some mechanism is needed to extract some particular item from the composite—Hintzman (1986, 1988) discusses several possibilities, such as comparing the echo to the stored traces, or repeating the retrieval process several times, each time using the retrieved echo as a probe, until the echo achieves a stable value (usually matching some stored trace). In any event, one basic difference between SAM and MINERVA 2 is that the latter model assumes that in recall a kind of composite memory trace is retrieved (at least initially), whereas the SAM model for recall holds that a specific memory trace is sampled (initially, though different traces may be sampled subsequently).

Distributed Memory Models

In recent years, composite/distributed memory models have enjoyed a rapidly growing popularity. [For additional discussion we refer the reader to a recent

Annual Review chapter by Hintzman (1990)]. These models fall in two related but somewhat different classes. In one class, items are represented by vectors (as in MINERVA 2) or matrixes of elementary features and the memory consists of a sum of the vectors or matrixes [e.g. TODAM (Murdock 1982), CHARM (Eich 1982; 1985), James Anderson's vector model (Anderson 1973), and the Matrix model (Pike 1984; Humphreys et al 1989); Kanerva's SDM model (Kanerva 1988) falls part way between this class and the separate storage class of the previous section]. In the second class, memory consists of nodes connected by weighted links; items are represented by a pattern or set of activations of the nodes, and long-term memory consists of the values of the weights on the links [e.g. Grossberg's ART model (Grossberg 1987; Grossberg & Stone 1986), James Anderson's BSB model (J. A. Anderson et al 1977), McClelland & Rumelhart's recurrent model, or any of the feedforward back-propagation models].

The basic difference between such models and the models discussed previously is that composite/distributed memory models assume that a memory trace is not a distinct, localized entity but rather part of a combination or superimposition of all traces input to the system. It is this aspect that has made many of these models seem both mysterious to the novice (who wonders how memory can be as good as it is) and attractive to many experts (who can explain why memory is as bad as it is, and how we can extract averages and prototypes from inputs, and who like the analogy to neuronal structures).

These composite/distributed storage assumptions can serve as a basis for a memory model because for each version there exists an appropriate retrieval operation. In some cases the cue will retrieve a noisy version of the original trace containing that cue; in other cases the cue will retrieve a noisy version of an item originally stored as an associate of the cue; in yet other cases the retrieval may be a clearly definable response, but with a type of noise determining the probability of reaching such a state, and determining whether the state would be the correct one. The retrieved information can be matched against the input to perform recognition, or if necessary can be "cleaned up" in some fashion to allow a response to be emitted.

As an example, consider one version of the Matrix model proposed by Anderson et al (1977; termed BSB for "brain state in a box"). Whenever two items (\mathbf{f}_i , \mathbf{g}_i) are associated, a matrix \mathbf{A}_i is produced with cell elements $\mathbf{A}_i(r,s) = \mathbf{f}_i(r)\mathbf{g}_i(s)$. The composite memory (\mathbf{M}) consists of the sum of all such association matrixes, $\mathbf{M} = \Sigma \mathbf{A}_i$. Ignoring for simplicity the details of the node activation process (such as its nonlinear limitations on activation growth), the retrieval of an associate (\mathbf{g}_i) given a cue item (\mathbf{f}_i) can be obtained by postmultiplying \mathbf{M} with \mathbf{f}_i ; the result, $\mathbf{M}\mathbf{f}_i$, is a noisy composite of those vectors that had been studied with both \mathbf{f}_i and items similar to \mathbf{f}_i (see Anderson et al 1977:417). Although this model is formulated specifically for paired-

associate recall, a simple modification will handle recognition memory. The basic idea is that recognition involves a matching operation of the composite memory trace with the to-be-recognized item. In order for this to work, it must be assumed that the memory trace includes not only associative information but also item information.

Although we have ignored the short-term activation features of the BSB model, it is a member of our second class. We begin our discussion of the first class with a model closely related to Anderson's but without a node activation process (and its nonlinear limit on activation values). This is the Matrix model proposed by Pike (1984; see also Humphreys et al 1989a,b). Associations of item vectors are represented by matrixes to form a composite memory matrix. In recent versions of this model, context-to-item associations have been incorporated in the matrix model in order to account for the fact that memory of a to-be-memorized list is to some extent "isolated" from all other memories. Thus, instead of storing a two-way association between the members of a paired associate, a three-way association among the two items and the context is stored (in the form of a 3-dimensional matrix). In order to retrieve g_i the memory matrix (\mathbf{M}) is multiplied in a specific way (see Humphreys et al 1989a) with the matrix obtained by multiplication of the context (\mathbf{x}) and item (\mathbf{f}_i) vectors. The latter product defines the "interactive" retrieval cue representing the association of context and stimulus item. This incorporation of contextual associations makes it possible to distinguish between episodic (list-specific) and semantic (preexisting) associations.

Related models have been proposed by Murdock (1982) and Eich (1982; 1985; see also Metcalfe & Murdock 1981). In both Murdock's Theory of Distributed Associative Memory (TODAM) and Metcalfe's Composite Holographic Associative Recall Model (CHARM),¹ the associative encoding and retrieval operations are the mathematical operations of convolution and correlation, respectively (see Eich 1982, 1985).

The TODAM model assumes that when each association $A-B$ is studied, the vectors representing A and B , and the convolution vector representing $A-B$, are all added to a slightly decayed version of the single composite memory vector that contains all of episodic memory. In this model, recognition involves matching the to-be-recognized item vector to the memory vector (i.e. taking the dot product) and using the resulting scalar number as a measure of familiarity. Recall starts by correlating the cue item vector with the memory trace, producing a noisy vector containing components representing versions of all items associated to the cue vector during study. The noisy vector must

¹The term "holographic" refers to the analogy between the properties of human associative memory and those of holograms (Pribram et al 1974; Willshaw 1981), in particular their resistance to local damage and the associative properties.

then be cleaned up to produce a response (say, by comparing the retrieved vector to a list of separately stored vectors representing items in semantic memory).

Rather than store vectors, CHARM (Eich 1985) stores autoconvolution vectors for each single item; these are stored along with the convolution vector for the association. For all studied pairs, these vectors are summed into a single composite memory vector. The retrieval operation is correlation, as in TODAM. In CHARM, a probe with *A* retrieves a composite noisy version of all items convolved with *A*, including *A* itself, so that item and associative information are not independently retrieved (as they are in TODAM). As usual, the trace must be cleaned up to generate a response in a recall task. Recognition can be accomplished by comparing the retrieved vector to the test vector.

The second class of composite/distributed model explicitly incorporates processes of node activation (often thought of as short-term memory) as well as weight modification (the set of weights representing long-term memory), both processes typically being nonlinear. The complexities introduced have led most investigators to explore these models in the form of computer simulations (with the notable exception of James Anderson and Steve Grossberg — see below). Such models are often described by the terms “connectionist” or “neural net.” Most applications have been to learning phenomena, categorization and classification, or perceptual phenomena, but some discussion of applications to memory is useful.

Consider first a representative back-propagation model (Ackley et al 1985; Rumelhart et al 1986). This model assumes a 3-layer representation: a layer of input units or network nodes, a layer of output units, and a middle layer of so-called hidden units. Activation is fed from the input units to the hidden units (using a nonlinear transform) and from these to the output units. All connections between layers have weights that determine how much the activation of a particular, say, hidden unit depends on the activation of a particular input unit. The basic rule of the back-propagation model is that these weights are adjusted during training in order to optimize the correspondence between predicted and actual output vectors (the back-propagation algorithm performs a kind of least-squares fitting procedure). One can use such a model to perform recognition and recall in a number of ways; perhaps the simplest is to have each input association attempt to reproduce itself at the output layer. Then a subsequent test with an item will tend to produce a noisy version of the association containing that item at the output layer. Recognition can be accomplished by matching, and recall by cleaning up the trace in some fashion.

These networks can represent virtually any computable mapping from input to output layer (given enough hidden units). However, for our purposes the

important issue is how such a mapping is learned and retained under various conditions. These aspects have been considered by McCloskey & Cohen (1989) and Ratcliff (1990).

McCloskey & Cohen (1989) showed that in two-list recall tasks the back-propagation model suffers from “catastrophic forgetting”: The second list leads to almost complete forgetting of the first list. Similar problems were uncovered by Ratcliff (1990), who analyzed the model’s predictions for recognition memory. This result is understandable if it is realized that these neural net models adjust the connection weights to fit the most recent stimuli, and that it is assumed that the inputs during the second list are of second-list items only. At the start of second-list learning, the weights will be configured optimally for the first list. However, there is no mechanism in the model that will keep the weights from obtaining completely different values, optimizing the “recall” of the second-list items. Hence, after a few training trials on the second list, the network will have “forgotten” the first-list items. Ratcliff (1990) also showed that this model fails to predict a positive effect of amount of learning on the d' measure for recognition. [Ratcliff also showed that several related models, including the auto-associative model proposed by McClelland & Rumelhart (1985)—see below, failed to resolve the problems. Research going on at the time of this writing suggests a number of new approaches that might work; e.g. Sloman & Rumelhart (1992); Kruschke (1992); Lewandowski (1992). Below we discuss Grossberg’s ART model, which deals with the problem explicitly.]

The back-propagation models are “feedforward” networks: Activation flows only forward through the system (the amount of error is in a sense propagated backwards through the system in order to adjust the weights appropriately, but this should not be confused with the flow of activation). On the other hand, a number of models are recurrent: Activation that leaves a node can be fed back to that same node, possibly after flowing through a number of intermediate nodes, and the process typically continues until a stable pattern of activation results. (The BSB model has this character, though we did not discuss the dynamics of activation.)

Consider first the McClelland & Rumelhart (1985) model. In brief, a set of nodes accepts input from external sources and is fully interconnected (except that nodes do not directly activate themselves) by directional links having weights. Activation moves through the system driven by the sum of the external and internal inputs to a node, until a stable pattern is reached. Then weights are adjusted so as to reduce the difference between the internal and external input to each cell (so that the network will try to reproduce its external inputs). Recognition can be accomplished by matching an input to the stable pattern of activation it produces, and recall by cleaning up the same stable pattern in some fashion.

Finally, consider the recurrent ART models of Grossberg (e.g. Grossberg & Stone 1986, although the original models date back before 1970). We describe here a greatly simplified version of the theory to give the flavor of the approach. Memory consists of a series of ordered layers of nodes. Consider just two layers, with perceptual inputs to layer 1, (and, in general, top-down inputs to layer 2). In addition, weighted links exist in both directions between nodes in the two layers, and activations pass in both directions along these links. Within a layer, there may also be connections, but these are inhibitory and do not carry activations directly. The two layers pass activation rapidly back and forth until a stable result is achieved. Because of inhibition within layer 2, a single node will come to be active in this layer, in stable resonance with a pattern of activation on the nodes in layer 1. The stable pattern may be used for recognition or recall in ways similar to those we have discussed already.

A particularly noteworthy feature of the model is its method for picking the single active node in layer 2. The pattern of activations sent down from this node to layer 1 is compared with the pattern in layer 1. If these do not match well, the currently active layer 2 node is turned off, the system resets, and a new node in layer two wins the competition. This continues until a good match is found, until a node not yet used as a template for a pattern is found, or until no nodes are left, in which case all layer 2 nodes become inactive. The result is that different patterns are assigned new nodes, and new learning does not harm old learning in the destructive fashion of other models of this class.

The weights on the links change continuously also, but at a much slower rate than the activation changes. The upward weight changes are made so as to reduce the difference between a weight itself and the signal passed upward along that link. Thus a set of weights leading to a single active node comes to correlate with the activation pattern in the nodes. Also the downward links from the active node are adjusted to match the activation pattern in the layer 1 nodes, so that top-down templates of the presented patterns are learned. The weights leading to and from any one layer 2 node come to encode a set of highly similar patterns, so that each node in layer 2 can be thought of as a category prototype. A particularly noteworthy feature of this system is the fact that the system can have a distributed representation at some levels (e.g. level 1) and a potentially separate representation at other levels (e.g. layer 2).

Differences and Similarities

In this section we compare the various models on a number of important theoretical dimensions. Although the various approaches we have considered are superficially quite different, basic phenomena are often explained in a similar manner, albeit using different terminologies. Here we focus on the basic issues concerning the conceptualization of memory processes. The

discussion deals with issues as if they were independent, but it is important to remember that no one of the hypotheses discussed below can be right or wrong in isolation; each must be analyzed in the context of the model in which it is embedded.

SEPARATE VS COMPOSITE MEMORY TRACES The question here is whether the model assumes that different items are stored in separate traces or in one composite memory trace. Whether this is a meaningful theoretical distinction depends on a number of auxiliary assumptions in the models in question. For example, Shiffrin & Murnane (1991) showed that an arbitrary number of events can be stored in a single number on a single link in a way that allows each event to be retrieved without error. The method is not a physically realizable one, however. Plausible composite systems, incorporating the equivalent of neural noise, all seem to have at least one testable property: When the system is densely composite, then the storage of new inputs, or even the repetitions of old inputs, tends to degrade the representations of other old inputs. Ratcliff et al (1990) tested this notion empirically and found that repetitions of some list items did not reduce recognition performance for other list items (see also, Murnane & Shiffrin 1991).

Shiffrin et al (1990) looked at the implications for extant models. All then-current models were found wanting. They concluded that composite models dense enough to predict forgetting caused by the composition property could not predict the findings. They concluded that models positing separate traces had the potential to predict the results, and they developed a variant of the SAM model that did so. This variant assumed that repetitions were accumulated in a single trace (a kind of local composition hypothesis—see below). It also incorporated a “differentiation” hypothesis: Suppose two different items *A* and *B* were not rehearsed together. If *B* is stored in memory more strongly, then *A* used as a cue will tend to activate it less.

A more local composition issue concerns whether two separate presentations of a given item are encoded separately or in the same trace. That is, if an item is repeated, does the second presentation lead to a strengthening of the originally formed trace, or will a new trace be formed?

MINERVA 2 assumes that each separate encoding of a single item (repetition) leads to a separate episodic memory trace, ACT* assumes that repetitions strengthen a given trace, and the early versions of SAM were somewhat ambiguous about this point. Recently (see Raaijmakers 1991), the SAM model has been extended to deal specifically with the effects of repetition and the spacing of repetitions. In this version, a kind of study-phase retrieval assumption has been added to the model. That is, on the second presentation an (implicit) retrieval attempt occurs. If the trace representing the first presentation is retrieved, it is assumed that the new information will be added to the

“old” memory trace; otherwise, a new trace will be formed. In addition, Shiffrin et al (1990) had to assume a strengthened trace to explain the lack of a list-strength finding in recognition.

This view gained support from a study by Murnane & Shiffrin (1991). They tried to induce separate storage of a repeated word by embedding it in different sentence contexts; this manipulation produced the expected positive list-strength effect in recognition. Evidence of a different sort supporting this view arises from a study by Ross & Landauer (1978). They showed that the traditional spacing effect only occurs for the probability of recalling single items presented twice and not for the probability of recalling either one of two different items each presented once. This result seems to require that repetitions of an item should be treated differently from the case of multiple items, each presented once (although firm conclusions depend on the details of each model).

REPRESENTATIONAL ISSUES We consider three representational aspects: 1. the nature of the information encoded in the memory trace, 2. whether links between memory traces are assumed, and 3. the representation of “associative strength.”

The models we have considered differ in their assumptions about the information encoded in the memory trace. In the all-or-none activation model ACTE of Anderson (1976), storage of a simple pairwise association involves the formation of a new link between pre-existing network nodes. In the ACT* model, what is stored is a cognitive unit representing the episodic experience. It is assumed that such a new network node has associative links with nodes representing the constituent parts of the item—i.e. (in this case) stimulus, response, and list context. In ACT*, associative strength is represented simply by the strength of the memory traces. As described above, these strengths determine the amount of activation that spreads to the trace from associated nodes.

SAM and MINERVA 2 also assume that the trace represents the “episodic experience” but are less specific about the exact nature of what is stored. The original SAM model focused on the relation between cues and images: Associative relations are represented by a “retrieval structure” rather than the more traditional “storage structure.” The model does not make use of explicit associative connections between images, though these are present implicitly in the following sense: Suppose two items are studied together; when one is used as a cue the retrieval strength to the image of the other is high. SAM was not entirely explicit concerning the nature of the “image,” though for most verbal studies an image was based on the individual word. Shiffrin et al (1989) presented evidence that a good deal more flexibility is needed, and that a sentence is often a single image (and that, under some circumstances, a pair

of words is a single image). Thus in principle a pair association could be stored in two ways: separate images governed by an implicit association that is represented in the retrieval structure, or a single combined image.

MINERVA 2 assumes that episodic experiences and memory traces can be represented as vectors of feature values. Since the nature of these features is left unspecified, this assumption does not really pose any restrictions. It is noteworthy that MINERVA represents pairs by encoding the component vectors back to back in a single stored vector. There is a nonlinear activation process during retrieval that lets the system distinguish whether two stored items are in the same or different traces.

Finally, what about the models with composite/distributed representations? The final representation is generally a vector or matrix; it is a composite of similar vectors or matrixes (or degraded forms of these) stored for individual items and pairs. The question is how associations and single items are handled during storage. TODAM has item vectors, and convolutions of item vectors for associations. Context information could in principle be part of each vector but in recent work has been treated as a separate vector. CHARM treats single items as autoconvolutions but is otherwise similar to TODAM in most respects. The Matrix model treats individual item vectors separately, and context as a separate vector. Single items are stored as an association matrix made from the item, context, and a unit vector. Pairs are matrixes made from the product of the two item vectors and the context vector. One issue left unresolved for these models is the basis on which some types of information are encoded in a given vector while other types are singled out for treatment as a separate vector. For example, how would category information be treated? (See Humphreys et al 1991 for one possible solution.)

A more general solution to this problem is possible if the various types of information, and various items to be associated, are all treated as components of a single vector, or single pattern of activation across a set of nodes. For example, in the McClelland & Rumelhart autoassociative recurrent model, and in Grossberg's ART models, all items to be associated, and related information, are encoded as a single vector or pattern of activation values sent to a set of nodes. Anderson's BSB model, and various versions of feedforward back-propagation models, use either of two methods. In one method, similar to those in the recurrent models just mentioned, items to be associated are encoded together in a single input vector [for example, the model of Ackley et al (1985) tries to reproduce at the output layer the vector presented to the input layer]. All such models use a pattern-completion property to retrieve associates. In the second method, the items to be associated are treated as separate vectors; for example, the input layer could encode one item and the output layer could encode the associated item (J. A. Anderson et al 1977).

CONTEXTUAL ENCODING Any model that is designed to explain data from episodic memory experiments must somehow account for the fact that a paired associate item such as *apple-engine* can be learned despite the presence of strong competitive semantic associations (*apple-pear*). It seems highly unlikely that one or two presentations of a list would create such a strong association that it dominates the pre-experimental associations. Thus, subjects are able to learn different associative relationships in different situations. This contextual dependence is a fundamental property of episodic memory.

It therefore seems highly desirable for a model of memory to have some means of selectively accessing memory traces stored under particular temporal or contextual conditions. Note that a simple recency-based mechanism will not suffice—subjects can access information from contextually identifiable periods in the past. To give just one example, subjects are able to access selectively not the list learned most recently but the list learned prior to that (Shiffrin 1970; Anderson & Bower 1972).

Most models incorporate contextual associations as the means to focus retrieval processes in episodic memory, either by including the contextual information in the memory trace or by treating the contextual information as a separate item. Whether this alone will suffice is an open question. For example, ACT* and MINERVA 2 assume an additive rule for combining the associative strength due to context and item. Such a rule may not have a sufficiently strong focusing effect to eliminate strong interference by pre-experimental associations. The multiplicative combination rules used, for example, in SAM and the Matrix model are such that retrieval is focused on those traces (or those components of the composite trace) that are consistent with the context at test. Even a multiplicative rule may not, by itself, be sufficient to focus retrieval properly. For example, Humphreys et al (1989a) call attention to crossed-associates lists, in which the subject is asked to learn pairs like *doctor-king* and *queen-nurse*. Versions of the SAM model in which individual words (but not pairs) have a single (semantic) memory representation would not easily predict the learning seen in such cases. However, SAM models typically assume that images are episodic in nature, not semantic.

It should be no surprise that models that do not incorporate context will not fare well. For example, a model that does not include a way to reduce the effect of irrelevant associations will have serious problems explaining why the interfering effect of the number of items on a single experimental list is not completely swamped by the millions of previously acquired associations. A simple forgetting assumption, e.g. a reduction of strength for previous associations (as in TODAM), will not do the job without added assumptions about context: The strong empirical list-length effects would require too rapid and massive forgetting.

STORAGE The issue we focus on here is the predicted effect of increasing study time for a list. TODAM, CHARM, and the Matrix model provide examples of models in which simply adding more copies of each trace to memory may not improve memory: Both the mean and standard deviation of the retrieved signal rise together in such a way that performance does not change (Murdock & Lamon 1988; Shiffrin et al 1990). At least two approaches have been used to solve this problem. Hintzman (1986, 1988) and Murdock & Lamon (1988; see also Murdock 1989) have proposed a probabilistic encoding assumption: Each feature of an item is encoded (stored) with a probability p that rises with presentation time. If not stored it is given a neutral value (or in a variant discussed by Shiffrin et al 1990, replaced by a random value). Eich (1985:28) proposes a variant in which all features of an item are either encoded or not (all-or-none encoding). In all these variants, both repetitions of an item and increased study time will improve storage relative to variance in the system, and therefore increase performance.

Shiffrin et al (1990) discuss an alternative way in which these models might show a learning or repetition effect. This alternative is based on the fact that performance in these models is related to the signal-to-noise ratio (or d'). Since d' measures the ratio of mean signal strength to the standard deviation, d' can show an increase with repetition if a constant is added to the standard deviation. The reason for this is that the standard deviation will no longer be completely proportional to the mean signal strength. It is natural to suppose that the constant represents activation of traces or trace components from lists other than the one being tested, or from extra-experimental memory. (More generally, this assumption may prove useful in all models because it lessens the effect of list variables like list length and study time in accord with the amount of extra-list activation).

The remaining models predict performance increases with repetitions or study time for fairly obvious reasons: storage of stronger associations in SAM or ACT*, or weight changes that produce better encoding in the neural net models.

RETRIEVAL One of the major differences among the models discussed here concerns the manner in which the retrieval process produces a recalled item. In SAM separate traces are accessed separately, so the recovered information can be compared to a standard lexicon; SAM doesn't provide any details of this process but simply assumes the probability of successful recall rises with the strength of the cues to image relationship. The ACT models use similar probabilistic rules. MINERVA 2 also has separate storage but retrieves a composite. This composite could be compared with the individual stored traces, but this seems unsatisfactory because recognition is also assumed to be

a composite process. Hintzman (e.g. 1988) proposes a somewhat more satisfactory solution in which the composite retrieved vector is used as a subsequent retrieval cue, the process continuing in this way until the retrieved vector comes to represent an unambiguous item. ART also has separate storage, and test probes come to activate some single node in at least one layer; this node sends down a pattern of activation that could itself produce a clear recall, or (if it is a category node rather than a single item node) could be compared to separately encoded patterns elsewhere in the system.

The models that assume composite storage and retrieval face a greater problem: How is the noisy retrieved trace cleaned up to allow an unambiguous recall? TODAM, CHARM, and the Matrix model assume a lexicon of separately stored items to which the retrieved trace can be compared. This solution tends to dilute the composite character of these models. The remaining connectionist and neural net models do not offer clear solutions for cases in which the retrieved trace is noisy enough to be ambiguous. Typically a probabilistic recall rule is adopted, based on the match of the retrieved trace to possible responses. If the model is fully composite, however, it is not entirely clear where the comparison stimuli lie.

A second issue involves whether the retrieval process is assumed to be probabilistic or not. Both ACT* and SAM assume a probabilistic retrieval process. In these models, an item that was not retrieved on a first retrieval attempt may still be retrieved if an additional attempt at retrieval is made. (In SAM it is usually assumed, however, that at least one new cue must be used for a subsequent retrieval to have a chance at success.) The other models, on the other hand, are such that a second attempt will always lead to the same result (unless the cues are changed, or have added noise (see McClelland 1991).

Finally, only a few models (namely SAM and the convolution/correlation model of Metcalfe & Murdock 1981) have been applied to extended search processes as in free recall, in which the subject uses a number of different retrieval cues in order to maximize recall. It might be argued that the search strategies that are probably involved in these paradigms are not part of the "basic" or "elementary" memory processes. However, such a viewpoint does not do justice to the fact that many real-life situations do involve this type of unstructured memory retrieval.

FORGETTING Let us define forgetting as a failure to retrieve information from memory at time *B* when it was retrievable at an earlier time *A*, or as a decrease in the probability of retrieval. There seem to be three basic ways in which forgetting might occur: 1. a decrease in the "strength" of the memory trace—i.e. decay; 2. an increase in competition by other, interfering, traces (or items); and 3. a change in the nature of the cue between time *A* and time

B—i.e. a change in the (functional) stimulus. There does not seem to be any difference between the models with respect to the third aspect, although not all of them have explicitly dealt with it. Mensink & Raaijmakers (1988) have used this idea in their application of the SAM model to interference and forgetting. In this model, part of the forgetting was assumed to be caused by contextual changes—i.e. changes in the contextual cues between study and test.

MINERVA 2 (Hintzman 1988:532) and TODAM (Murdock 1989:77) both assume that trace information is subject to decay. In TODAM, this is built into the basic equation of the system (see Murdock 1982, Eq. 1). It should be noted, however, that this decay assumption differs somewhat from traditional decay conceptions in that it is related to the storage of new information in the memory trace: Each time a new item is added to the composite memory trace, a fixed proportion of what was there is lost, producing a strong recency effect with a geometric character. (It should be noted that this assumed decay is independent of the interference that is common to all of the composite storage models, including TODAM.) All the memory models considered in this chapter predict a decrease in performance due to learning other items, and to learning other pairs of items (in both *AB-AC* and *AB-CD* type tasks), the only general exception occurring when the other items are rehearsed or coded jointly with the items in question. In general, several mechanisms in each model help produce interference; these mechanisms may be different for different tasks (as in SAM), and the mechanisms may differ between models. We mention here a few of the more interesting differences among the models.

Most composite models incorporate explicit interference due to the superimposed storage assumptions. When vectors or matrixes are added together, or when a set of weights are jointly adjusted for each new input, the result tends to be degradation of the representations of each item. There are of course exceptions to this rule: If memory is large enough relative to the size of the inputs, then storage might be effectively separate (the amount of superimposition might be minimal; see Kanerva 1988), or if the inputs are orthogonal enough, or if the system orthogonalizes or separates the inputs (e.g. Grossberg's ART models), then interference would not be mandated by the factor of composite storage.

The remaining sources of forgetting are posited to arise during the course of retrieval (in SAM these are the only sources of forgetting). SAM assumes summation of activations at retrieval to accomplish recognition; as a consequence, extra items cause forgetting by increasing "noise." In MINERVA, composition during both recall and recognition causes interference due to increasing noise. One chief remaining cause of interference is based on the relative strength of storage of different items. For example, in SAM, sampling in recall is based on a ratio of activation strengths. Reduction in relative

strengths of targets due to extra items also plays an important role in many of the models under discussion, especially the ACT models. This factor plays a chief role in accounting for list-length, fan, and cue-overload effects.

APPLICATIONS TO PARADIGMS

In this section, we compare the ways the various models predict certain basic findings in memory research, both qualitative and quantitative.

Cued Recall

Cued recall is the basic paradigm for associative memory, and the present set of models have been formulated so that cued-recall predictions can be made. ACT* and CHARM have been applied more or less exclusively to cued-recall data. In addition, ACT* and SAM have been shown to be able to account for both latency as well as accuracy data in cued recall (Anderson 1981; Mensink & Raaijmakers 1988).

List Length

All the models are capable of handling the basic list-length effect. However, in some models (TODAM, CHARM) no distinction is made between (*a*) the to-be-recalled list and (*b*) extra-list and extra-experimental information. When list length is predicted to have an effect, it does so because retrieval is restricted to the to-be-recalled list (without explanation). This seems unsatisfactory, and the natural way to resolve the difficulty would be the adoption of some form of contextual cuing (as is the case with other models).

However, whether a contextual cue is used may be less important than how it is used. A typical multiplicative rule for cue combination tends to focus access upon regions of memory in the intersection of the sets of memory traces evoked by each cue separately, whereas a typical additive rule tends to access traces in the union of these sets. Humphreys et al (1991) argue convincingly for the intersection approach, implying that “strengths” or “activations” should be acted upon in a way functionally equivalent to multiplication (as in the SAM model, the Matrix model, etc) rather than addition (as in ACT*).

This type of explanation of list-length effects sees such effects as an example of a more general effect—i.e. that the efficacy of any probe cue is inversely related to the number of memory traces or items associated to that probe cue (which might be called the length of the list of associated items).

Interference and Forgetting

The basic issues here are the effects of different types of interference (i.e. AB-AC vs AB-CD), mechanisms for (relative) spontaneous recovery, single-

list forgetting paradigms, and whether or not some sort of decay notion is used. ACT*, SAM, and CHARM have all been explicitly applied to such phenomena.

Anderson (1983a) and Mensink & Raaijmakers (1988) show that some results in this area necessitate the assumption that recall is based on both relative and absolute associative strengths. Relative strength is a function of the number and strength of other associations, while absolute strength is indexed by the amount of study time or the number of presentations of an item. In ACT*, absolute strength comes in through the assumption that trace formation is more likely as the total study time increases. In SAM, absolute strengths determine the probability of recovering enough information from the trace to give the name of the item as a response.

Mensink & Raaijmakers (1988) present a theoretical analysis of traditional interference phenomena. They show that modern memory models such as SAM can reconcile phenomena that have been problematic for traditional interference theories. Such analyses bring out a number of tacit assumptions in the typical verbal (i.e. nonquantitative) models that are not usually noted.

Free Recall

This paradigm is more complex than cued recall. This is due to the fact that it necessitates not only an exact formulation of the relation between STS/working memory and long-term memory but also a description of search/retrieval strategies. Only a few of the models have dealt explicitly with such data. We briefly discuss predictions by SAM (Raaijmakers & Shiffrin 1980) and an early version of the CHARM model (Metcalf & Murdock 1981).

SAM assumes that contextual and inter-item associations are built up as a result of rehearsing the items in STS. A buffer process (Atkinson & Shiffrin 1968) is used to model the rehearsal process. Retrieval starts by outputting any items still in STS. Thereafter the retrieval process is modeled as a series of retrieval attempts either with the context cue alone or using both context and a previously retrieved item as probe cues. This process continues until the number of failed searches reaches a specific criterion.

One of the strong points of the SAM model is that it handles with a single set of parameter values data from lists with large variations in presentation rate and list length. The latter result is predicted because the search termination criterion is exceeded sooner for the longer lists, relative to list length: Relatively fewer samples are made from a longer list than from a shorter list. This prediction is characteristic of sampling-with-replacement search models with a fixed stop criterion. It also subsumes the cue-overload principle proposed by Watkins (1975; see also Mueller & Watkins 1977; Watkins & Watkins 1976). This principle states that the probability of recalling any

particular item decreases with the number of instances associated to the retrieval cue.

Although the Metcalfe & Murdock (1981) free-recall model is described in the terminology of the convolution-correlation model, the actual simulation model does not in fact use the mathematical operations of convolution and correlation. Instead, all-or-none associations are stored between list items, and between list items and context (treated as a list item). When an item is used as a cue, a random choice is retrieved perfectly from the stored associates, if any.

The rehearsal process is conceptualized as a continuous cuing of the memory vector with the currently available item. Thus, when an item is presented, it is associated to the item that is currently available (context at the start of the list). Then the just-presented item is used as a cue to generate an item to which it has been associated, and then this item is used as a new cue, etc. This continues until the next item is presented, which is then associated to the item that is currently available.

At the time of recall, the last rehearsed item is recalled and used as a cue to generate another item; then this item is itself used as a cue, and this continues until a certain criterion period has passed without any new items being recalled. At that point, context is reinstated as a cue and the process begins anew and continues until the criterion period passes for the second time without any new recalls.

Each of these models predicts serial position effects. Since the SAM model is based on the two-store framework, it should not come as a surprise that it makes many of the same predictions as the classic two-store model (see Atkinson & Shiffrin 1971; Raaijmakers & Shiffrin 1980), and for the same reasons: Primacy is predicted because of the cumulative rehearsal assumption, while the output from STS leads to a recency effect. Although the two-store model is often described in textbooks as having problems handling data on levels-of-processing and recency effects, this is in fact not correct (see Raaijmakers 1991).

The Metcalfe-Murdock model has a quite different flavor. In this model, the shape of the serial-position curve is critically determined by the cues available at recall. Recency is predicted because the last-presented item is recalled first and then used as a cue. This item is assumed to be the optimum entry point into the end of the list. The disappearance of the recency effect by the introduction of a delay between presentation and test is explained by the assumption that rehearsal continues during the delay.

Hence, at the end of the delay the currently available item will most likely be some item other than the last item on the list. The optimum entry point for recall of the last few items is therefore lost. This explanation seems unlikely since providing the subject with the terminal item after the delay interval

should reinstate the recency effect. (Another problematic aspect is the assumption that rehearsal continues during the delay filled with arithmetic.)

Primacy is predicted by this model because context is used as a retrieval cue (in the second phase of the recall process), and context is nearly always relatively strongly associated with the first item. Thus, this explanation is quite similar to the typical two-store explanation of primacy as being due to stronger traces for the initial items (in this case being more strongly associated to context).

One of the more important advantages of the recent work on models of memory is that it has led to model-based simulation programs for specific experimental tasks. These programs can then be used to see how the model behaves under specific experimental conditions. This is especially important in free recall since this paradigm does not lend itself easily to analytic approaches.

One aspect of the data where this has been proven helpful is in the analysis of the effects of various types of cuing manipulations of the likelihood of recall. We mention two: the (positive) effects of category cues and the (negative) effects of cuing with randomly selected list items (the so-called part-list cuing effect).

Raaijmakers & Shiffrin (1980) showed that typical effects of cuing with category names could be easily predicted by the SAM simulation model. These predictions do not greatly depend on the specific assumptions of SAM (*vis-à-vis* alternative models). Such analyses are, however, important to show that observed effects are indeed consistent with particular theoretical frameworks.

This is even more the case in the part-list cuing paradigm. In this paradigm subjects are given some randomly selected items from the list as cues for the remaining list items. The typical finding is that such cuing leads to a slight but unexpected decrease in the probability of recall for the remaining items. Raaijmakers & Shiffrin (1981) spent a good deal of effort analyzing this peculiar effect within their SAM simulation of free recall. They showed that this counterintuitive effect was in fact predicted by the model. In addition to the basic result, a number of related findings were predicted. These included the effect of the number of cues, the time at which the cues were given, and the effect of interpolated learning (between presentation and test). Raaijmakers (1991) shows that the model predicts a reversal of the cuing effect if a delayed testing procedure is used. This prediction is indeed borne out. This research has also shown that it is by no means easy to intuit the predictions of a relatively simple model such as SAM in a complicated experimental situation.

This part-list cuing effect has also been dealt with by Metcalfe & Murdock (1981). However, in their simulation it was assumed that the list cues were

not actually used by the subject. This assumption makes it relatively easy to predict a negative effect of cuing but does not make much sense given the fact that most subjects will expect the cues to be helpful (as did most memory specialists). In addition, such an approach makes it impossible to predict a reversal of the cuing effect in delayed cuing.

Note, however, that these problems are not due to the basic structure of the convolution/correlation model (which was in fact put aside) but to the assumptions that are made concerning the subject's strategy. This illustrates that predictions for free-recall tasks depend critically on the strategy used—i.e. on the assumptions made concerning the sequence of cues to be used in retrieval. Gronlund & Shiffrin (1986) examined the effects of various retrieval strategies on recall from natural categories and categorized lists. They showed that different strategies indeed have an effect on recall performance. This result poses two problems for any model for free recall. First, it makes it problematic to apply a specific (arbitrary) version of the model to the data of a group of subjects, unless it can be shown that the result of interest is insensitive to the choice of strategy or that the subjects all use a similar strategy. Second, given a specification of retrieval strategies (i.e. in terms of the sequence of cues that are used), the model should be able to give a quantitative account of the resulting performance differences. Gronlund & Shiffrin (1986) show that a simple extension of SAM could account for the observed differences.

Recognition

Most current models of memory assume that simple recognition decisions are based on some sort of global familiarity value. By this we mean that the familiarity value is a kind of weighted, additive combination of the activation of all items in memory. This global familiarity value is determined by the match between the probe cues and the memory trace(s). This general type of model has been termed the General Global Matching Model (GGMM, Humphreys et al 1989b) or the Interactive Cue Global Matching (ICGM) model (Clark & Shiffrin, submitted). As these labels imply, such models differ from previous local matching models in that all items in memory are involved in the match, not just the representation of the tested item. In this section we consider some of the data used to test these models.

Pair Recognition

Pair recognition has been used as an experimental paradigm to test aspects of recognition models. Basically, the issue here is the way associative information is assumed to contribute to recognition decisions. In these experiments the subject first studies a list of word pairs (A_1B_1, A_2B_2, \dots). At test, intact pairs (A_iB_i) have to be discriminated from rearranged pairs (A_iB_j), mixed pairs

(A_iX), and/or new word pairs (XY). These results may be compared to those obtained in single item recognition (A_i vs X) and/or cued recognition (A_iB_i vs A_iX where only the second item has to be judged; see Clark & Shiffrin submitted). Humphreys et al (1989b) show that all extant versions of the global matching model (SAM, MINERVA 2, Matrix, and TODAM) lead to similar equations for the mean matching strengths. This would seem to imply that it will be difficult to differentiate between these models. However, predictions for d' depend not only on the mean strengths but also on the variances. Furthermore, it may be possible to distinguish between the models if one also takes the predictions for single-item recognition and cued recognition into account.

Clark & Shiffrin (submitted) examined the predictions for all types of recognition tests. They show that the models differ with respect to whether they predict an advantage for cued recognition compared to single-item recognition. The results of their experiments were reasonably well predicted by TODAM and SAM, with TODAM producing the best fit. MINERVA 2 and the Matrix model did not fit the data well. One problem with such data, however, is that it might very well be the case that subjects make use of recall processes in addition to global matching. That is, the logic of the models allows subjects to supplement global matching with recall.

Gronlund & Ratcliff (1989) pointed to another problem for global matching models. They examined the time course of the availability of item and associative information using a response-signal procedure (Reed 1973, 1976; Doshier 1976). In this procedure, a recognition decision must be made at one of several predefined times after the onset of the test stimulus. With this procedure it is possible to determine the growth of accuracy as a function of processing time. Gronlund & Ratcliff showed that item information becomes available sooner than associative information. This poses a problem for global matching models since these treat these two types of information as inseparable. To accommodate the results, separate contributions of item and associative information are required, possibly by distinguishing between concurrent and compound usage of cues (see Gronlund & Ratcliff 1989). That is, it might be assumed that memory is probed in parallel with an interactive, compound cue and with the item cues separately. As an alternative, it might be the case that pair images are sometimes stored, and that the time course of pair-image activation differs from that of single-item image activation.

List Length vs List Strength

Recent research by Ratcliff et al (1990) has focused on the effects of the strength of other list items on the recall and recognition of target items. This so-called "list-strength effect" concerns the effects of strengthening (or

weakening) some list items upon memory for other list items. Ratcliff et al (1990) showed that strengthening some items in the list decreases recall of the remaining list items but has no or even a positive effect on recognition performance. This contrasts with the list-length effect: Adding items to a list decreases both recall and recognition performance. Thus, the number of irrelevant items, but not their strength, affects recognition. This is true not only for strength variations due to amount of study time but also for variations due to spaced repetitions.

This peculiar result should have a number of consequences for models of recognition. In particular, it will be necessary to assume a structural difference between presentation of two different items and two presentations of a single item. Shiffrin et al (1990) showed that current memory models indeed cannot predict both the presence of a list-length effect and the absence (or reversal) of a list-strength effect.

Shiffrin et al (1990) also investigated whether the various models could be modified to enable prediction of these results. Such modification does not seem possible for models that assume items are stored in one composite memory trace. Even considering recognition only, these models cannot predict both a positive list-length effect and an absent or negative list-strength effect when strength variations are due to spaced repetitions. Models such as SAM and MINERVA 2 that assume separate storage are in principle better equipped to handle these results, although they too will have to be modified to enable prediction of negative list-strength effects.

Shiffrin et al (1990) show that a modification of SAM can handle these results. In this modified SAM model it is assumed that different items are stored in separate traces but repetitions of an item within a list are stored in a single memory trace. Second, the variance of activation of each separate trace, when the cue item is unrelated to the item(s) encoded in the trace, is constant regardless of the strength of the trace. The latter assumption is inconsistent with previous formulations of SAM but is defended using a differentiation argument: The better the image is encoded, the clearer are the differences between it and the test item, and hence the lower the activation. In this way, a constant or even decreasing variance may be predicted, depending on the weighting of context and item cues.

A crucial aspect of this explanation is that repetitions of an item are assumed to be stored in a single memory image. To evaluate it further, Murnane & Shiffrin (1991) tested whether a reversal of the list-strength effect in recognition occurs if repetitions are presented in such a way that they are likely to be encoded in separate images. They found that repetitions of words in different sentences produced a list-strength effect whereas repetitions of entire sentences did not. This demonstrates that the nature of the encoding of a repeated item is a crucial factor.

Evaluating the Models

In this chapter we have shown that current mathematical models of memory are capable of handling many classical and new findings in recall and recognition. We suggest that models of this type are superior to verbally stated theories of memory. Arguments in favor of the modeling approach include these: 1. the ability to predict the size (and not just the direction) of the effect of experimental factors, 2. the ability to predict the effect of combinations of experimental factors, 3. the ability to examine the combined result of theoretical assumptions, 4. the fact that a model (especially in the form of a simulation program) can be used to “experiment” with processing or strategy assumptions to determine the crucial variables that underlie a given prediction, and 5. the fact that models often demonstrate the limitations of more intuitive reasoning.

Such a conclusion is, however, often criticized on the ground that general models of memory of the type discussed in this chapter are too versatile. That is, the models usually incorporate a relatively large number of processes and parameters that seem to enable them to predict almost any type of empirical result. In addition, it is often difficult to intuit what a specific model will predict in a given situation. This contrasts with the simplicity of typical verbal, nonquantitative explanations of memory phenomena. In this section we argue that this difficulty is often more apparent than real.

First, quantitative models also make qualitative predictions that do not depend on parameter values. That is, in order to evaluate a model’s ability to predict data, one should not only examine the phenomena that it can predict but also take into account whether it makes strong, parameter-free predictions about results that should not occur (no matter what parameter values are used). Second, if a particular prediction depends on the specific parameter values used, it should be possible to arrange the experimental situation in such a way that that particular result is reversed. Third, the argument may also be turned around: If the ability of a model to predict a particular phenomenon turns out to depend on parameter values one may well ask whether a corresponding qualitative explanation is in fact logically sufficient. Finally, some results are indeed complex (i.e. dependent on a number of interacting processes) whether we like it or not. In fact, one advantage of quantitative models of the type discussed in this chapter is that they may be used to see whether particular “verbal” explanations hold true when tested in the context of a comprehensive model of human memory. The next sections focus on specific aspects of this discussion.

Number of Parameters

Current quantitative models of memory frequently incorporate a dozen or so parameters. These parameters reflect both structural aspects of the memory

system (decay rates, processing times) and task-related aspects (weighting of cues, stopping criteria, decision criteria). When a model is fitted to a set of data, these parameters usually have to be estimated from those data; that is, they are given values so as to optimize the fit to the data. Although this procedure can be rigorously defended on statistical grounds, it does seem to many to involve a bit of cheating. It is probably for this reason that the relatively large number of parameters in current models is frowned upon.

In many cases however, the number of parameters is not really an issue. That is, the qualitative nature of the predictions does not depend on the exact parameter values. Thus, many of the simulations are performed using a single set of parameter values (see, for example, Metcalfe & Murdock 1981; Mensesink & Raaijmakers 1988; Hintzman 1988). In those cases where parameter values do reverse a particular prediction, empirical evidence should be obtainable concerning this prediction (see, for example, the prediction of a reversal of the part-list cuing effect as a function of the contextual strength parameter in the SAM model; Raaijmakers 1991). Another point is that nonquantitative models also include parameters—that is, degrees of freedom—although this is rarely realized. To put this in another way, most explanations for memory phenomena by models of memory might be formulated in a qualitative way. In this way, there would not be any basic difference between quantitative and qualitative models. However, the resulting theories would have lost most of their explanatory power.

Number of Processes

Most of the difficulties with well-specified quantitative models have to do with the relatively large number of processes that are usually proposed. This is especially the case when models attempt to be applicable to a large number of different experimental paradigms. As emphasized by Smith (1978), there is a tradeoff between generality and simplicity of theoretical models. The problem here is that due to the number of processes and the number of parameters (or quantitative relations) involved in complex memory models, it is often not possible to make predictions about the behavior of the model except through quantitative simulations.

An example (drawn from own experience) illustrates this point. When the SAM model was first applied to the part-list cuing paradigm (see Raaijmakers & Shiffrin 1981), it was not at all clear whether it would or would not predict this effect. Furthermore, even after the prediction turned out to be successful, it was not immediately clear (to say the least) what factors in the model were causing it.

What this shows is that it is not possible to make intuitive predictions about the behavior of a model under specific task conditions. However, it should be evident that a similar problem holds for “verbal” theories of memory. In such

qualitative accounts it is not clear what the boundary conditions are that apply to a particular prediction. The lesson that can be drawn here is that much more effort should be invested in theoretical analyses of the factors involved in predicting empirical phenomena. Such analyses should focus on the role of each of the proposed processes in the explanation of a particular phenomenon.

Quantitative, Qualitative, and Semi-Quantitative Fits

Although all the models that we have considered are formulated in a quantitative manner, there is a tendency in current work to restrict the analysis to qualitative predictions; that is, one analyzes only whether a model predicts the general direction of an effect, rather than the exact magnitude. In contrast to the tradition of the 1960s and 1970s, typical goodness-of-fit measures such as the chi-square statistic do not figure prominently in the typical article that nowadays presents a formal model of memory.

This poses a problem. On the one hand, it can be defended that one does not want to focus too specifically on the exact numerical details of one particular experiment; on the other hand, it would be desirable at least to look at the relative magnitude of a particular effect (relative to other predicted effects). That is, suppose that there are two phenomena of interest, effect *A* and effect *B*, where *A* is a large effect and *B* a small (but consistent) one. It is conceivable that a model would be able to predict both *A* and *B* in a qualitative manner but that it would always predict either *A* and *B* both small or *A* and *B* both large. Such a "misfit" would not be detected if the analysis focuses only on the qualitative aspects.

Fortunately, most presentations of formal models of memory employ a strategy that falls between these two extremes. The typical approach is to use a single set of parameters to examine a set of data (or data patterns) that is representative of empirical findings. Although none of the actual data are really fitted in the traditional sense, the use of a single set of parameters makes it possible to verify that the model makes predictions in the right ballpark in terms of relative effect sizes.

Hence, we may distinguish among three degrees of comparing the model to actual data: qualitative, quantitative, and what might be called semi-quantitative analysis. The first involves only the direction of a difference between conditions; the second involves a direct comparison between the predicted and observed data using a goodness-of-fit measure; finally, the third does not involve a goodness-of-fit measure but does look at the sizes of the predicted and observed effects.

Although real quantitative fits remain a desirable feature, it might be argued that the proper approach is to aim first for a semi-quantitative prediction of the data. In this phase, the emphasis is on showing that a model can deal with a variety of findings from different task paradigms. At some point, a

number of promising models will have been developed. At that stage, the time seems to be ripe for quantitative tests in which several models may be compared in terms of goodness-of-fit. We believe that the demonstrated potential of current models of memory justifies the expectation that future work in this area will involve more comparative, quantitative testing.

Quite recently, John Anderson (1990; see also Anderson & Milson 1989) has proposed a model that attempts, in a sense, to meld some of the best features of the two approaches we have been contrasting (detailed, formal, quantitative, process models vs general, verbal, descriptive models). His "Rational" model bypasses details of representation and process to the greatest possible degree, and instead is aimed at the general proposition that memory is organized so as to solve the memorizer's problems in an optimum fashion. In any given retrieval situation, it is assumed that each event stored in memory has a number assigned to it representing its probability of being relevant (containing the desired information). It is assumed that these events are searched in order of their relevance, either until a retrieval occurs or a stopping criterion is reached. The probabilities are based on two multiplicative factors: the past history of an event's usefulness (independent of the cues used to probe memory) and the likelihoods of relevance associated with the cues. So far only the barest hints of applications to memory paradigms are available. It is interesting that the model operates at a very abstract level and yet offers quantitative predictions for certain phenomena. Although initial results are intriguing, it is far too early to assess the long run usefulness of the approach.

ACKNOWLEDGMENTS

This research was supported in part by Grants NIMH 12717 and AFOSR 870089 to the second author.

Literature Cited

- Ackley, D. H., Hinton, G. E., Sejnowski, T. J. 1985. A learning algorithm for Boltzmann machines. *Cogn. Sci.* 9:147-69
- Anderson, J. A. 1973. A theory for the recognition of items from short memorized lists. *Psychol. Rev.* 86:417-38
- 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. *Psychol. Rev.* 84:413-51
- Anderson, J. R. 1976. *Language, Memory, and Thought*. Hillsdale, NJ: Erlbaum
- Anderson, J. R. 1981. Interference: the relationship between response latency and response accuracy. *J. Exp. Psychol.: Hum. Learn. Mem.* 7:326-43
- Anderson, J. R. 1983a. A spreading activation theory of memory. *J. Verb. Learn. Verb. Behav.* 22:261-95
- Anderson, J. R. 1983b. *The Architecture of Cognition*. Cambridge, MA: Harvard Univ. Press
- Anderson, J. R. 1990. *The Adaptive Character of Thought*. Hillsdale, NJ: Erlbaum
- Anderson, J. R., Bower, G. H. 1972. Recognition and retrieval processes in free recall. *Psychol. Rev.* 79:97-123
- Anderson, J. R., Milson, R. 1989. Human memory: an adaptive perspective. *Psychol. Rev.* 96:703-19
- Atkinson, R. C., Shiffrin, R. M. 1968. Human memory: a proposed system and its control processes. In *The Psychology of*

- Learning and Motivation: Advances in Research and Theory*, ed. K. W. Spence, J. T. Spence, 2:89-195. New York: Academic
- Atkinson, R. C., Shiffrin, R. M. 1971. The control of short-term memory. *Sci. Am.* 224:82-90
- Bower, G. H. 1961. Application of a model to paired-associate learning. *Psychometrika* 26:255-80
- Clark, S. E., Shiffrin, R. M. 1992. Associations, retrieval capacity, and cued recognition. *J. Exp. Psychol.: Learn. Mem. Cogn.* submitted
- Doshier, B. A. 1976. The retrieval of sentences from memory: a speed-accuracy study. *Mem. Cogn.* 8:291-310
- Eich, J. M. 1982. A composite holographic associative recall model. *Psychol. Rev.* 89:627-61
- Eich, J. M. 1985. Levels of processing, encoding specificity, elaboration, and CHARM. *Psychol. Rev.* 92:1-38
- Estes, W. K. 1955. Statistical theory of spontaneous recovery and regression. *Psychol. Rev.* 62:145-54
- Gillund, G., Shiffrin, R. M. 1984. A retrieval model for both recognition and recall. *Psychol. Rev.* 91:1-67
- Gronlund, S. D., Ratcliff, R. 1989. The time-course of item and associative information: implications for global memory models. *J. Exp. Psychol.: Learn. Mem. Cogn.* 15:846-58
- Gronlund, S. D., Shiffrin, R. M. 1986. Retrieval strategies in recall of natural categories and categorized lists. *J. Exp. Psychol.: Learn. Mem. Cogn.* 12:550-61
- Grossberg, S. 1987. Competitive learning: from interactive activation to adaptive resonance. *Cogn. Sci.* 11:23-63
- Grossberg, S., Stone, G. 1986. Neural dynamics of word recognition and recall: attentional priming, learning, and resonance. *Psychol. Rev.* 93:46-74
- Hintzman, D. L. 1984. MINERVA 2: a simulation model of human memory. *Behav. Res. Methods; Instrum. Comput.* 16:96-101
- Hintzman, D. L. 1986. "Schema abstraction" in a multiple-trace memory model. *Psychol. Rev.* 93:411-28
- Hintzman, D. L. 1988. Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychol. Rev.* 95:528-51
- Hintzman, D. L. 1990. Human learning and memory: connections and dissociations. *Annu. Rev. Psychol.* 41:109-39
- Humphreys, M. S., Bain, J. D., Pike, R. 1989a. Different ways to cue a coherent memory system: a theory for episodic, semantic, and procedural tasks. *Psychol. Rev.* 96:208-33
- Humphreys, M. S., Pike, R., Bain, J. D., Tehan, G. 1989b. Global matching: a comparison of the SAM, MINERVA II, Matrix, and TODAM models. *J. Math. Psychol.* 33:36-67
- Humphreys, M. S., Wiles, J., Bain, J. D. 1991. Memory retrieval with two cues: Think of intersecting sets. In *Attention and Performance XIV: A Silver Jubilee*, ed. D. E. Meyer, S. Kornblum. Hillsdale, NJ: Erlbaum
- Kanerva, P. 1988. *Sparse Distributed Memory*. Cambridge, MA: MIT Press
- Kruschke, J. 1992. ALCOVE: an exemplar-based connectionist model of category learning. *Psychol. Rev.* in press
- Lewandowski, S. 1991. Gradual unlearning and catastrophic interference: a comparison of distributed architectures. In *Relating Theory and Data: Essays on Human Memory and Learning in Honor of Bennet B. Murdock*, ed. W. E. Hockley, S. Lewandowski, pp. 445-76. Hillsdale, NJ: Erlbaum
- McClelland, J. L. 1991. Stochastic interactive processes and the effect of context on perception. *Cogn. Psychol.* 23:1-44
- McClelland, J. L., Rumelhart, D. E. 1985. Distributed memory and the representation of general and specific information. *J. Exp. Psychol.: Gen.* 114:159-88
- McCloskey, M., Cohen, N. J. 1989. Catastrophic interference in connectionist networks: the sequential learning problem. *Psychol. Learn. Motiv.* 24:109-65
- Mensink, G. J., Raaijmakers, J. G. W. 1988. A model for interference and forgetting. *Psychol. Rev.* 95:434-55
- Mensink, G. J. M., Raaijmakers, J. G. W. 1989. A model of contextual fluctuation. *J. Math. Psychol.* 33:172-86
- Metcalfe, J., Murdock, B. B. 1981. An encoding and retrieval model for single-trial free recall. *J. Verb. Learn. Verb. Behav.* 20:161-89
- Mueller, C. W., Watkins, M. J. 1977. Inhibition from part-set cuing: a cue-overload interpretation. *J. Verb. Learn. Verb. Behav.* 16:699-709
- Murdock, B. B. Jr. 1982. A theory for the storage and retrieval of item and associative information. *Psychol. Rev.* 89:609-26
- Murdock, B. B. Jr. 1989. Learning in a distributed memory model. In *Current Issues in Cognitive Processes: The Tulane Flowerree Symposium on Cognition*, ed. C. Izawa, pp. 69-106. Hillsdale, NJ: Erlbaum
- Murdock, B., Lamon, M. 1988. The replacement effect: repeating some items while replacing others. *Mem. Cogn.* 16:91-101
- Murnane, K., Shiffrin, R. M. 1991. Word repetitions in sentence recognition. *Mem. Cogn.* 19:119-30
- Murnane, K., Shiffrin, R. M. 1992. In-

- terference and the representation of events in memory. *J. Exp. Psychol.: Learn. Mem. Cogn.* In press
- Pike, R. 1984. A comparison of convolution and matrix distributed memory systems. *Psychol. Rev.* 91:281-94
- Pribram, K., Nuwer, M., Baron, R. 1974. The holographic hypothesis of memory structure in brain function and perception. In *Contemporary Developments in Mathematical Psychology*, ed. R. C. Atkinson, D. Krantz, R. D. Luce, P. Suppes, pp. 416-57. San Francisco: Freeman
- Raaijmakers, J. G. W. 1991. The story of the two-store model: past criticisms, current status, and future directions. See Humphreys et al 1991. In press
- Raaijmakers, J. G. W., Shiffrin, R. M. 1980. SAM: A theory of probabilistic search of associative memory. *Psychol. Learn. Motiv.: Adv. Res. Theory* 14:207-62
- Raaijmakers, J. G. W., Shiffrin, R. M. 1981. Search of associative memory. *Psychol. Rev.* 88:93-134
- Ratcliff, R. 1990. Connectionist models of recognition memory: constraints imposed by learning and forgetting functions. *Psychol. Rev.* 97:285-308
- Ratcliff, R., Clark, S., Shiffrin, R. M. 1990. The list-strength effect. I. Data and discussion. *J. Exp. Psychol.: Learn. Mem. Cogn.* 16:163-78
- Reed, A. V. 1973. Speed-accuracy trade-off in recognition memory. *Science* 181:574-76
- Reed, A. V. 1976. List length and the time course of recognition in immediate memory. *Mem. Cogn.* 4:16-30
- Ross, B. H., Landauer, T. K. 1978. Memory for at least one of two items: test and failure of several theories of spacing effects. *J. Verb. Learn. Verb. Behav.* 17:669-80
- Rumelhart, D. E., Hinton, G. E., Williams, R. J. 1986. Learning internal representations by error propagation. In *Parallel Distributed Processing: Explorations in the Microstructures of Cognition. Vol. 1: Foundations*, ed. D. E. Rumelhart, J. L. McClelland, pp. 216-71. Cambridge, MA: MIT Press
- Shiffrin, R. M. 1970. Forgetting: trace erosion or retrieval failure? *Science* 168:1601-3
- Shiffrin, R. M., Murnane, K. 1991. Composition, distribution, and interference in memory. See Lewandowski 1991, pp. 331-46
- Shiffrin, R. M., Murnane, K., Gronlund, S., Roth, M. 1989. On units of storage and retrieval. In *Current Issues in Cognitive Processes: The Tulane Flowerree Symposium on Cognition*, ed. C. Izawa, pp. 25-68. Hillsdale, NJ: Erlbaum
- Shiffrin, R. M., Ratcliff, R., Clark, S. 1990. The list-strength effect. II. Theoretical mechanisms. *J. Exp. Psychol.: Learn. Mem. Cogn.* 16:179-95
- Sloman, S., Rumelhart, D. E. 1992. Reducing interference in distributed memory through episodic gating. In *From Learning Theory to Cognitive Processes: Essays in Honor of William K. Estes*, ed. A. F. Healy, S. M. Kosslyn, R. M. Shiffrin. Hillsdale, NJ: Erlbaum
- Smith, E. E. 1978. Theories of semantic memory. In *Handbook of Learning and Cognitive Processes. Vol. 6.: Linguistic Functions in Cognitive Theory*, ed. W. K. Estes. Hillsdale, NJ: Erlbaum
- Watkins, M. J. 1975. Inhibition in recall with extralist "cues". *J. Verb. Learn. Verb. Behav.* 14:294-303
- Watkins, M. J., Watkins, O. C. 1976. Cue-overload theory and the method of interpolated attributes. *Bull. Psychonom. Soc.* 7:289-91
- Willshaw, D. 1981. Holography, associative memory, and inductive generalization. In *Parallel Models of Associative Memory.*, ed. G. E. Hinton, J. A. Anderson, pp. 33-104. Hillsdale, NJ: Erlbaum