Progress in Mathematical Psychology - 1
E.E. Roskam and R. Suck (Editors)

A FORMAL MODEL FOR ASSOCIATIVE MEMORY

Jeroen G.W. Raaijmakers

Department of Psychology, University of Nijmegen

Abstract

A review is given of the most important applications of the SAM theory for retrieval from long-term memory developed by Raaijmakers and Shiffrin (1981). It is shown how the general retrieval theory may be applied to paradigms such as free recall, cued recall, paired-associate recall, and recognition. It is emphasized that such across-task or across-paradigm analyses represent a major advance in mathematical modeling of learning and memory processes.

1 INTRODUCTION

There has been a long tradition of mathematical modeling of learning and memory processes. In recent years, however, there has been a marked change in the scope of mathematical models for human memory. In the sixties, the field was dominated by Markovian models. These models had a rather limited range of application, most of them being specifically designed for one particular experimental task. There were models for paired-associate memorizing, quite different models for recognition memory, and yet other models for free recall performance and short-term memory tasks, but none of these was applicable beyond those self-imposed boundaries. Although such models are still proposed from time to time and do serve important functions, a more recent trend is the development of what might be called 'general theories', instead of models developed for one specific experimental task. We now have what might be properly called 'models of memory' instead of, say, a model for forced-choice recognition tasks.

This paper will be about one such a general theory developed in the past few years by Richard Shiffrin of Indiana University and myself (Raaijmakers, 1979; Raaijmakers & Shiffrin, 1980, 1981a,b; Gillund & Shiffrin, 1984). This theory is denoted *SAM*, which stands for *Search of Associative Memory*. I will not go into all of the details of the particular models that have been developed for various applications, but, instead, I will try to show how a large number of empirical phenomena can be explained within the same theoretical framework. I will present applications of the theory to free recall, cued recall, recognition and paired-associate recall.

The general SAM theory assumes that long-term memory may be partitioned into complex information structures called 'images'. Such episodic images contain various types of information. We explicitly distinguish between (a) *contextual information* that can be used to identify the temporal-contextual setting in which the image was stored, (b) *item information* that can be used to reconstruct the item encoded in the image, and (c) *interitem information* encoding the associative relations between this image and other images.

SAM is based on the conception that memory is cue-dependent (Tulving, 1974), that is, what is elicited from memory is determined by the retrieval cues utilized at that moment. A number of images will be activated by the cues, in different degrees, depending on how strongly they are associated to the probe cues. In recall tasks in which the name of an item has to be recovered, one image is sampled from the activated set during each cycle of the memory search. In simple recognition, the decision is assumed to be based on the integrated activity over the entire activated set (Gillund & Shiffrin, 1984).

The starting point of any SAM analysis is the specification of a *retrieval structure* (see Figure 1). Such a retrieval structure gives the strength of relationship between each possible retrieval cue and each possible image. These retrieval strengths determine the probability that a given image will be retrieved when a particular set of probe cues is used. The entries in the retrieval structure are determined by (a) the coding and rehearsal processes used at the time of study, (b) preexperimental episodic images

retrieval cues	I_1	I_2	....	I_N
Q_1	$S(Q_1,I_1)$	$S(Q_1,I_2)$	....	$S(Q_1,I_N)$
Q_2	$S(Q_2,I_1)$	$S(Q_2,I_2)$	....	$S(Q_2,I_N)$
⋮	⋮	⋮		
Q_m	$S(Q_m,I_1)$	$S(Q_m,I_2)$	....	$S(Q_m,I_N)$

FIGURE 1

Retrieval structure used in SAM. This matrix gives the strengths that govern the retrieval process. Entries are the strengths from individual cues to individual images.

associations, and (c) the match of the cue encodings at study and test.

2 FREE AND CUED RECALL

I will begin by discussing the application of the theory to recall tasks, especially free recall. In such tasks, the retrieval process is divided into two phases, called *sampling* and *recovery*. It is assumed that memory search consists of a series of retrieval attempts, each consisting of a sampling and recovery phase. The probability of sampling a particular image is a function of the strength of association between the probe cues and that image. Specifically, the probability of sampling image I_i when a set of probe cues, Q_1, Q_2, .., Q_m, is used, is given by the following equation:

$$P_S(I_i|Q_1, Q_2, \ldots, Q_m) = \frac{\prod_{j=1}^{m} S(Q_j,I_i)^{W_j}}{\sum_{k=1}^{N} \prod_{j=1}^{m} S(Q_j,I_k)^{W_j}} \qquad [1]$$

This equation looks quite complicated. However, its meaning can be understood more easily if we define

$$A(i) = \prod_{j=1}^{m} S(Q_j, I_i)^{w_j} \quad [2]$$

as the activation of image i given the probe cues Q_1, Q_2, .., Q_m. The w_j in this equation are weights assigned to the different cues representing their relative saliency or importance. (In many of the applications it is unnecessary to assume unequal weights, and they are usually set equal to 1.0.) With this substitution, the sampling equation becomes a simple ratio rule:

$$P_S(i) = \frac{A(i)}{\Sigma\, A(k)} \quad [3]$$

Hence, the probability of sampling a given image is proportional to the activation of that image.

The key to the present approach is the product rule used to combine the individual cue strengths into a single activation. This multiplicative feature has the useful and important consequence that it allows focusing of the search. The images with the highest probability of being sampled are those with the highest product of strengths, and hence those that tend to be strongly associated to *all* of the cues. In other words, the sampled image tends to come from the most dense region of the intersection of the associative fields of the separate cues.

When an image is sampled, not all of its features will be activated. It is assumed that the proportion of image elements that will be recovered, is determined by the associative strengths between that image and the retrieval cues. In recall tasks where the name of the item encoded in the image is requested, we have used the following equation for the probability of successful recovery:

$$P_R(I_i|Q_1, Q_2, \ldots, Q_m) = 1 - exp\{- \sum_{j=1}^{m} w_j S(Q_j, I_i)\} \quad . \quad [4]$$

This rule transforms the summed associative strengths to a number between 0 and 1. This equation is not as arbitrary as it may look on first sight. It corresponds to the assumption that each cue gives an independent chance of successful recovery.

To apply this model to a particular experimental paradigm, one has to specify the storage assumptions and the retrieval strategy that is assumed. Such a retrieval strategy specifies the sequence of retrieval cues that are used by the subject. The storage assumptions and the retrieval strategy are of course dependent on the particular task and the instructions given to the subjects.

Let us now briefly describe the most important features of the free recall paradigm. In such a task, a list of n unrelated words is presented in a random order at a rate of t seconds per word. Both the list length and the presentation time may be varied between lists. Usually, the subject is asked to recall as many words as possible, in any order, immediately following list presentation. In some cases, an interpolated task, usually arithmetic, is given between presentation and test to eliminate effects due to short-term memory.

For the standard free recall task, it is assumed that the images in the retrieval structure can be limited to the items that were presented on the list. All other images are assumed not to be activated to a significant degree by the probe cues used. This assumption is acceptable since we assume that the list context is always part of the probe set, and, hence, the sampling process will be focused on the list items.

In the SAM model that we have used for free recall, the only cues considered are the general context cue, C, representing the list context, and the words from the presented list, W_1, W_2, .., W_n. A probe set always consists of context alone or of context along with one of the word cues. Hence, our storage assumptions should specify the strengths of the images to these probe cues.

It is assumed that these associative strengths are proportional to the length of time that a word is studied in short-term store. We have assumed a buffer model as a model for the rehearsal process in short-term store. The buffer size is r. New words enter the buffer until it is full; then each new word replaces a randomly chosen word already in the buffer. Thus, if we let t_i be the time that item i stays in the buffer, and t_{ij} be the time that the items i and j are together in the buffer, then we asume:

$$S(C,I_i) = at_i \ ,$$
$$S(W_j,I_i) = bt_{ij} \ ,$$
$$S(W_i,I_i) = ct_i \ , \qquad [5]$$

where a, b and c are parameters to be estimated. It should perhaps be mentioned explicitly that we do not assume that study time is the only factor affecting the storage strengths as some critics seem to believe (Bradley & Glenberg, 1983). Other factors such as the type of coding can be handled simply by letting the storage parameters depend on such factors. Finally, if two words on the list are not rehearsed together, a small residual strength value, d, is assumed. Thus,

$$S(W_j,I_i) = d \quad if \quad t_{ij}=0 \ . \qquad [6]$$

The most important part of the model consists of the retrieval process assumed. Figure 2 gives a flowchart for this retrieval process. At the beginning of test, any words still remaining in STS are output. Then retrieval from LTS begins. It is assumed that the subject stops retrieval once the total number of failures or unsuccessful retrieval attempts reaches a critical value, termed K_{max}. A failure is every retrieval attempt that does not result in the recall of a new word. It is also possible to use a consecutive failure rule, i.e. a rule that specifies that the subject stops once a given number of consecutive failures is reached. This has little effect on most of the quantitative predictions of the model.

Since in a standard free recall task no cues are given to the subject, the subject starts by using the context cue to sample from the images that are associated to that cue. This cue continues to be used until some new word is recalled. Whenever this happens, the next probe will consist of two cues, the context cue and the just recalled word. As can be seen from the flowchart, a given word cue will be used until L_{max} consecutive attempts with this cue set have failed. When this happens, the subject drops that word cue and returns to using only the context cue.

A final assumption that should be mentioned is that some learning occurs during retrieval. Whenever a successful recovery occurs, the strengths of

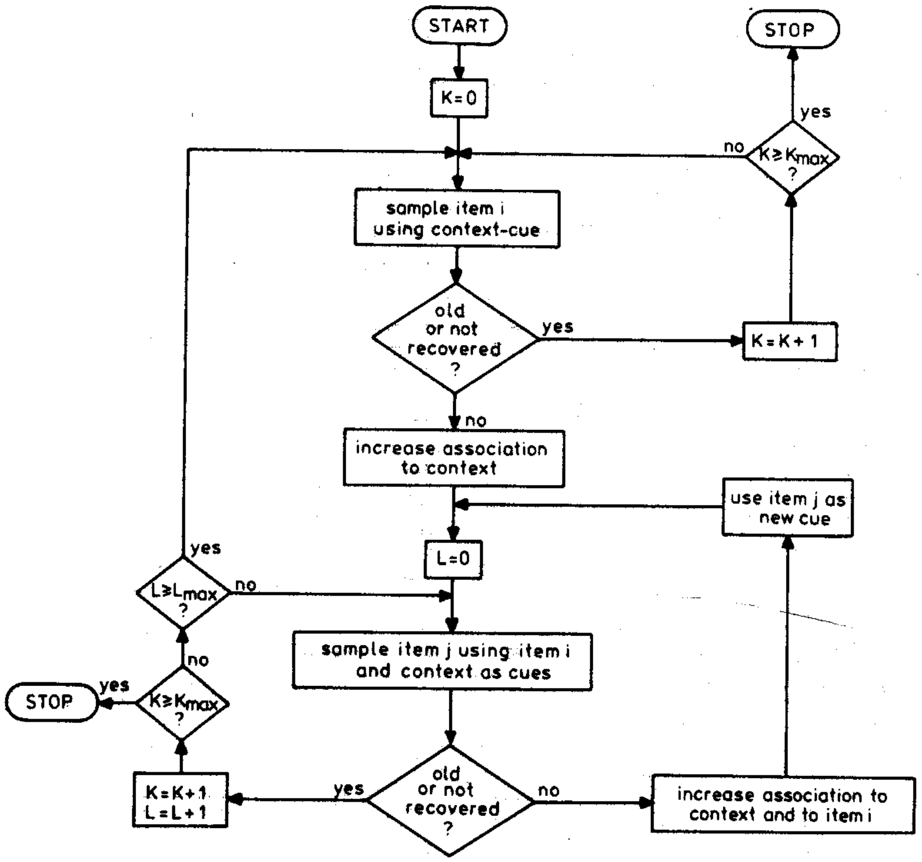

FIGURE 2

A flowchart of the main phase of retrieval in free recall. (From Raaijmakers & Shiffrin, 1980).

the probe cues to the recovered image are increased or incremented. This *incrementing* process obeys the following rules:

$$S'(C,I_i) = S(C,I_i) + e \quad ,$$
$$S'(W_i,I_j) = S'(W_j,I_i) = S(W_i,I_j) + f \quad ,$$
$$S'(W_i,I_i) = S(W_i,I_i) + g \quad . \qquad [7]$$

Here the primes represent the strengths after incrementing, and e, f, and g are the incrementing parameters. In most applications these parameters are all set equal to each other.

In summary, extensive use is made of interitem associative routes: whenever a new word is recalled it is used as a cue either until L_{max} failures accumulate or until a new word is recalled, in which case the new word is used as a cue. Of course, it could be argued that all interitem routes have not been fully explored, since a switch to a new cue may occur before the previous word cue has been exhausted. For this reason, a final *rechecking* process is incorporated in the model after the K_{max} criterion has been reached. In this phase, each previously recalled word is used as a cue along with context. L_{max} samples are made with each such probe set.

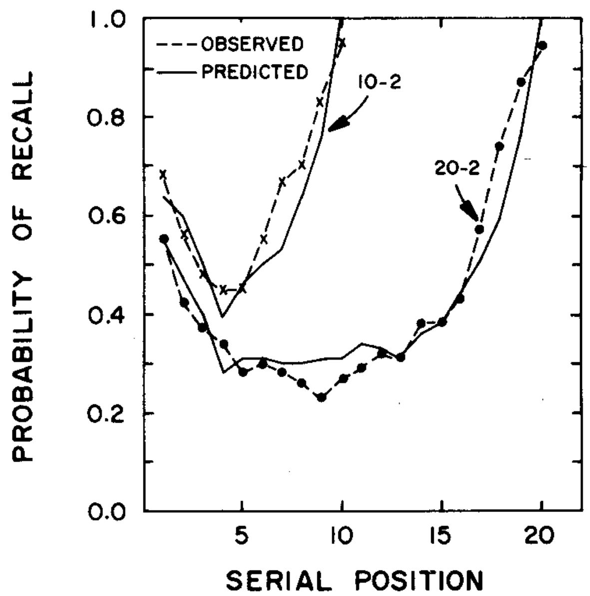

FIGURE 3

Serial position results and SAM's predictions for 10- and 20-word lists at 2 sec/word. (Data from Murdock, 1962). (From Raaijmakers & Shiffrin, 1980).

This rather simple model has been shown to be able to predict a large number of standard results in the free recall paradigm. Figure 3 shows observed and predicted serial position curves for two conditions from an experiment by Murdock (1962). The primacy and recency effects that are predicted by our model are a consequence of the buffer assumption and are not very informative concerning the LTS retrieval process. A more interesting result is that the model is able to predict the serial position curves for different list lengths and presentation times with the same set of parameter values.

Note that a list-length effect is predicted by the model: the probability of recall is a decreasing function of list length. This list-length effect is predicted by the model because *relatively* fewer samples are made from a longer list than from a shorter list. The probability of sampling an item is therefore lower for an item from a longer list. This phenomenon seems to be a quite general characteristic of retrieval processes: *the larger the number of items associated to a cue, the smaller the probability that any one of those items will be recalled*. This basic aspect of cue-dependent memory has been termed the *cue-overload principle* and has been used by Watkins to explain a number of empirical phenomena (see Mueller & Watkins, 1977; Watkins, 1975; Watkins & Watkins, 1976). Thus, it is of some interest to note that this cue-overload principle can be derived from the SAM theory.

A more extensive set of data was collected by Roberts (1972). Four list lengths (10, 20, 30, or 40 words) and five presentation rates (.5, 1, 2, 4, or 8 sec per word) were covaried. Figure 4 shows observed and predicted mean number of words recalled. It is evident that the qualitative features are predicted very well. The model correctly predicts that for a given list length the mean number of words recalled is a negatively accelerated function of the total presentation time. In addition, these data show that the total-time hypothesis (Murdock, 1960) is incorrect: equal total presentation times do not yield equal levels of recall.

The model also accounts for a number of time-dependent aspects. In free recall, the output rate decreases rapidly as more and more items are recalled. It has been known for some time that sampling models such as SAM give a good description of such interresponse times. In addition, the model predicts that recall will increase, albeit very slowly, if subjects could somehow be persuaded to continue searching even though they are practically not retrieving any new words. Such data have been collected by Roediger and Thorpe (1978) who managed to get their subjects to continue to try to recall for as long as 21 min. Such data can be predicted by the

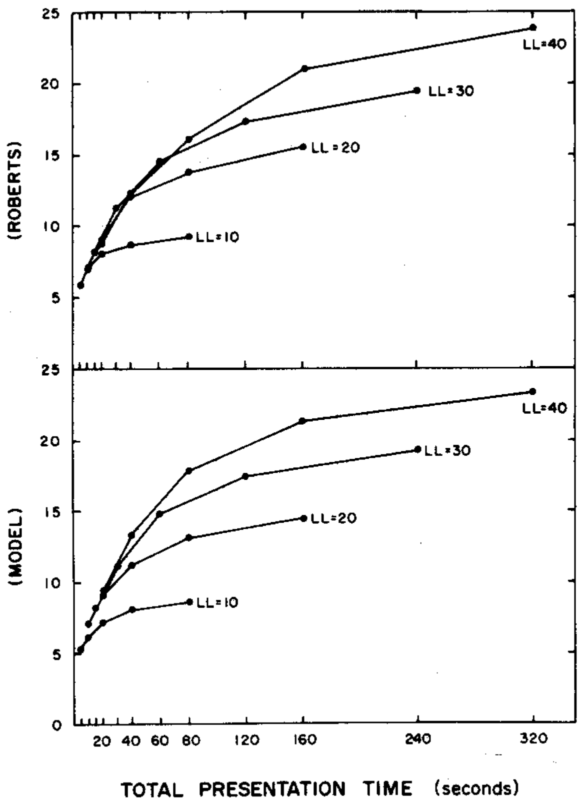

FIGURE 4

Observed data (Roberts, 1972 - top panel) and predictions of SAM (lower panel), for mean words recalled as a function of presentation time and list length (LL). (From Raaijmakers & Shiffrin, 1980).

SAM model simply by deleting the stop-rule. These results are related to an intriguing phenomenon called *hypermnesia*. This refers to the observation that when several consecutive recall periods are given, the total number of items recalled in these successive periods increases. Although this is not immediately evident, it turns out that the SAM model predicts such a hypermnesia result, especially with relatively long recall periods. This prediction is the result of two factors within the model: (a) the assumption that the associative strengths are incremented upon successful recall, and (b) the assumption that a previously sampled but not recovered image may still be recovered later if the cue set contains at least one cue that is new for that image. It may be shown that if both of these assumptions are deleted from the model, the hypermnesia result is no longer obtained.

Perhaps the single most important aspect of this SAM model for free recall is its prediction of the socalled *part-list cuing effect*. This effect is obtained when after list presentation a random subset of the list items is presented to the subjects in the experimental or cued group, who are told to use these words as cues to aid recall of the remaining list items, called the 'critical' or 'target' items. The control group is given no cues and recalls freely, as in the standard free recall task. On the assumption that these extra provided cues give the subject additional entry points into the associative memory network, one would have expected that the cued group would recall more critical items than the control group. The surprising finding, which was first obtained by Slamecka (1968), was that, in fact, the control group was slightly superior to the cued group. Slamecka and most other researchers in this area have argued that any theory that assumes that interitem associations are used in recall, should predict that at least some of these cues would facilitate recall of items that would otherwise not have been recalled. The fact that no recall facilitation has been found, has led these researchers to the conclusion that interitem or horizontal associations are not formed during study or at least not used during retrieval.

Application of the SAM model to this part-list cuing paradigm (see Raaijmakers & Shiffrin, 1981a) has however revealed that this reasoning is not correct. We have been able to show that a prediction of this effect is inherent in SAM-like models, despite the heavy use of interitem associative structure that is made in such models. In fact, it is this very structure and its use in retrieval that produces the effect. In Raaijmakers and Shiffrin (1981a) we have presented the results of a very extensive exploration of this issue. I shall now present only a brief summary of this work.

One of the most interesting results in this area has been the finding that the effect does not seem to depend very much on the similarity of the list items to each other, and hence on how easy interitem associations could have been formed. The results of an experiment by Slamecka (1968, Exp VI) show that there was no difference between the control and the cued groups, although the number of critical words recalled was doubled by the increasing similarity. Figure 5 shows the predictions of the SAM model for this type of experiment. It should perhaps be noted that this result presents a problem for the conventional explanation of the part-list cuing

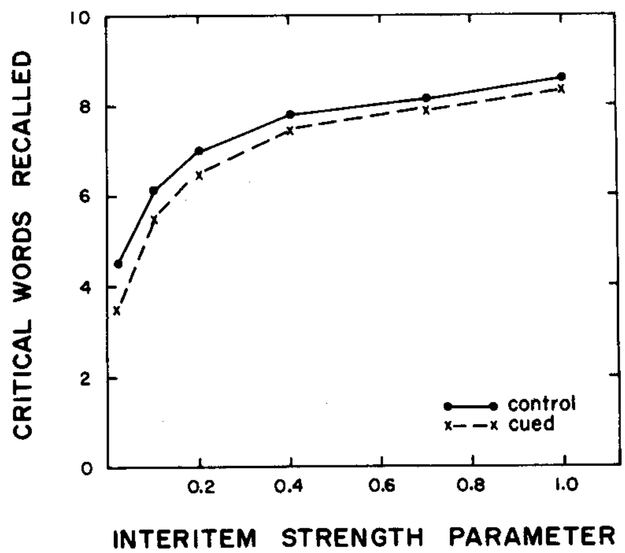

FIGURE 5

Predictions of the SAM model for the part-list cuing effect as the interitem strength parameter is varied. (From Raaijmakers & Shiffrin, 1980).

effect. That is, it is not easy to see how the increase in the number of words recalled with increasing associability should be explained without making use of any concept that resembles the use of horizontal associations. Our results show that the SAM model has no difficulty in predicting this result. In this application it was assumed that each of the cues provided by the experimenter was used by the subject along with the context cue until L_{max} failures are reached. A recovery of another cue word is not counted as a failure, except on the second and subsequent recoveries of the same cue word. This assumption makes the cue condition completely comparable to the control condition in the sense that the search process is equally effective in both conditions.

We have spent a considerable amount of time trying to find out why the SAM model predicts this part-list cuing effect, which aspects of the model are responsible for this prediction. It was found that the effect was not due to the assumption that the strengths of recovered items are incremented, making resampling more likely. Setting the increment parameters equal to 0 has an equal effect on both conditions. Another factor that is not responsible for SAM's prediction of this effect, is the particular stopping rule used nor the value for the stopping criterion (K_{max}) used. All of this is quite surprising since all of these predictions are made with a model that does include a factor that favours the cued group. That is, the cued group makes more item-plus-context searches, and the probability of recovery is higher when more cues are being used.

In order to understand why SAM predicts the part-list cuing effect it is necessary to realize that both conditions make extensive use of interitem searches. The only difference is the type of item cues used by the two groups. Subjects in the control condition make use of self-generated cues, whereas the cued group uses mostly experimenter-provided cues. Once this is realized, it is evident that any difference between the two conditions must be due to differential effectiveness of subject-generated versus experimenter-provided cues. In order to understand why self-generated cues are superior it is helpful to consider the following simplified associative network (see Figure 6).

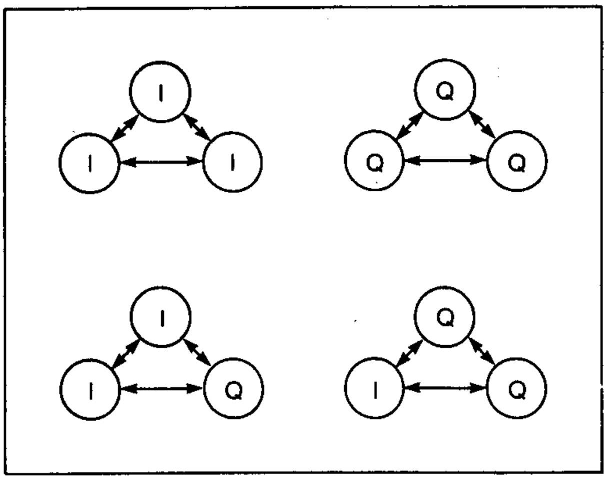

FIGURE 6

A simplified associative network that illustrates SAM's explanation for the part-list cuing effect. The cue items are denoted Q. The critical items are denoted I. See text for explanation.

Assume for the sake of argument that the words are interassociated in groups of three with all other interword associations being negligible. It is further assumed that sampling of an image from a triad leads to recall of all members of that triad. Assume furthermore that both groups sample an equal number of triads. It is important to note that the cues given to the cued group are a random sample from the items on the list. It is easy to see that the cued group will initially sample clusters of items which all contain at least one cue word. The experimenter-provided cues are therefore less efficient than the subject-generated cues because of a kind of artefact due to the scoring procedure: the subjects are only scored on the critical or target items, and the clusters sampled by the cued group will contain a relatively small number of target items.

This SAM model is able to explain most of the important results that have been obtained in this paradigm. Moreover, it is able to predict those conditions for which a positive cuing effect has been obtained. For example, the only paradigm for which a reasonably large positive cuing effect has been obtained, is a retroactive inhibition paradigm (Basden, 1973; Blake & Okada, 1973). It is reasonable to assume that in such a paradigm the context-to-item strengths are quite low at the time of testing. When we set the corresponding parameter in our model to a very low value, we indeed find a positive cuing effect. The reason for this is that in this case the control group is no longer able to generate enough cues by using the context cue. The model also predicts that there will be a slight positive cuing effect if we give the cues *after* a period of normal free recall instead of at the beginning of that period. Such a result has been found in an experiment by Allen (1969). Finally, the model predicts that the negative effect of list cues increases slightly with the number of such cues, a result which has also been found in a number of studies (see Roediger, 1974)

What is perhaps the most significant aspect of the model concerning these cuing results is that the model is able to predict with essentially the same mechanism both the negative part-list cuing effect as well as the large positive cuing effects that are obtained when the list is composed of a number of categories and the subject is given the category names as cues. For example, Figure 7 gives the observed and predicted number of words recalled for an experiment by Tulving and Pearlstone (1966) in which list length and number of items per category were varied. The subjects were given either a cued or a noncued (free) recall test. It is seen that cuing

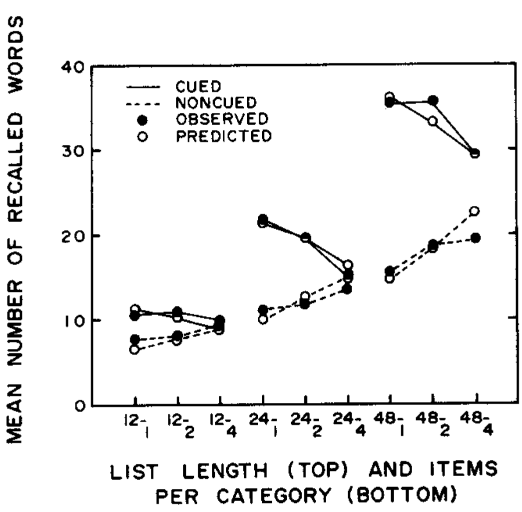

FIGURE 7

Observed and predicted mean number of words recalled as a function of list length and number of words per category for cued and noncued tests. (Data from Tulving & Pearlstone, 1966). (From Raaijmakers & Shiffrin, 1980).

has a large positive effect, especially if the list contains many different categories. Cuing has a positive effect in this case because the cues are optimally related to the associative structure: the subject is now given one good cue from each of the clusters. The cued group will then access more clusters than the noncued group. Cuing in this case alleviates the accessibility problems that the subject has in free recall. This is clearly shown in Figure 8 which gives for each condition the probability of recalling a category, that is, the probability of recalling at least one item from that category.

3 PAIRED-ASSOCIATE PARADIGMS

It might be noticed that the model developed for free recall can be easily extended to paired-associate recall paradigms. In fact, we have shown that free recall and paired-associate recall can be fitted simultaneously. This

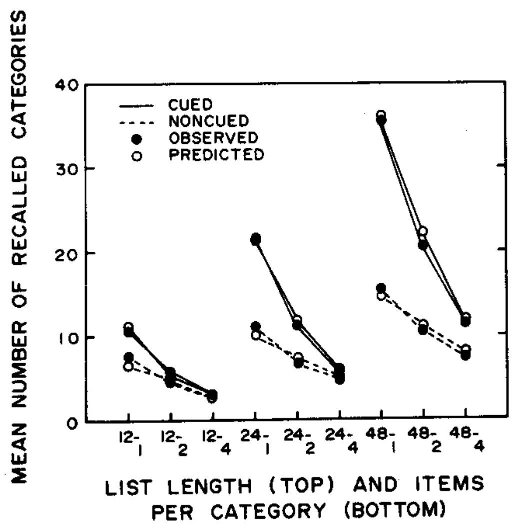

FIGURE 8

Observed and predicted mean number of categories recalled as a function of list length and number of words per category for cued and noncued tests. (Data from Tulving & Pearlstone, 1966). (From Raaijmakers & Shiffrin, 1980).

illustrates the advantage of a general theory for retrieval over a specific model that applies only to one particular experimental task. Starting from the general theory it becomes relatively easy to construct a model that applies simultaneously to tasks that have traditionally been treated separately and for which separate models have been proposed.

We ran an experiment (Raaijmakers & Shiffrin, 1981b) in which the lists were composed of pairs of words and single words. There were separate retention tests for these two types of items: a free recall test for the single words and a paired-associate recall test for the paired words. The main results are shown in Figure 9. The probability of recall on the paired-associate test decreases as a function of the number of pairs on the list and as a function of the number of single words on the list. Similar list-length effects are observed for the probability of recalling the single words on the free recall test. Note that these list-length effects take place even though recall is directed specifically to either the single words or the pairs. The fact that a list-length effect is also obtained in this kind of single-trial paired-associate recall procedure is quite

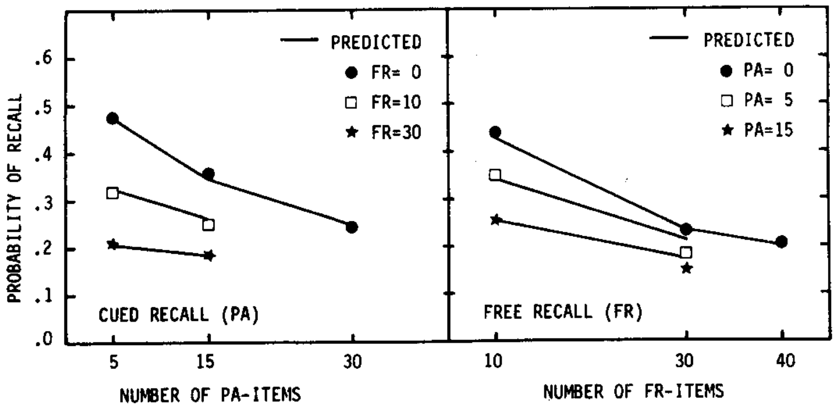

FIGURE 9

Observed and predicted probabilities of recall for single words and paired words as a function of the number of items of either type. (From Raaijmakers & Shiffrin, 1981b).

interesting since many models and theories for paired-associate recall do not predict such list-length effects. Figure 9 also shows the predictions from the SAM model developed for this paradigm, and it is evident that such effects are predicted by the SAM theory. The SAM theory predicts these effects because it assumes that the context cue is associated to all items on the list, and that a member of a pair is not only associated to the other member of that pair but also, very weakly, to all other images. Hence the presence of these other images has an interfering effect since there is a small probability that they will be sampled instead of the target image.

Another interesting aspect of these SAM models is that they predict output interference or test order effects in cued recall tasks. This effect has been observed for the first time in cued recall of categorized lists (see Smith, 1971; Roediger, 1973). In this paradigm the subject is presented a list composed of a number of categories, and at the time of recall, each category is tested in a successive manner by cuing with the category name. The output interference effect refers to the observation that the probability of recalling category members is highest for the first category tested and decreases in a systematic way for successively tested categories. The SAM theory predicts this effect because it assumes that the contextual associations are incremented upon recall. This increases the

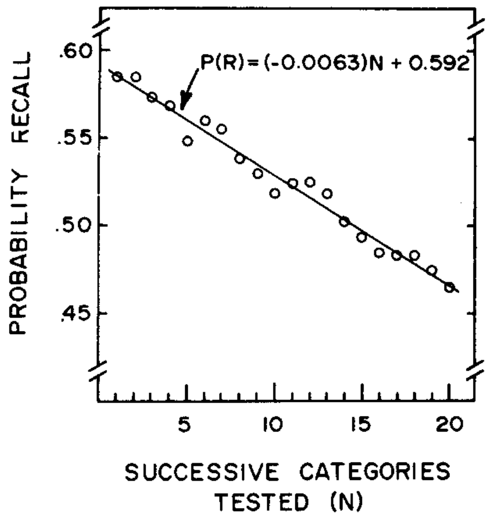

FIGURE 10

Predicted probability of recall as a function of test position of the category in cued recall. (From Raaijmakers & Shiffrin, 1980).

interfering effect of the categories tested earlier: there will be an increasing tendency to sample recalled items from the earlier categories, because their associations to the context cue have been incremented. This explanation is essentially the same as the explanation for the list-length effects in paired-associate recall discussed earlier. Figure 10 shows the predictions of the SAM theory. These predicted results were obtained with a model that had been successfully applied to fit the results of the experiment by Tulving and Pearlstone (1966) and the parameter values were identical to those used in that application. Note that this figure only gives the predicted results, based on a very large number of simulations. The observed results are much more variable, but the overall magnitude of the effect is predicted quite well by this SAM model. A similar effect has been observed in paired-associate recall (Roediger & Schmidt, 1980; Raaijmakers & Shiffrin, 1981b). Such retrieval inhibition effects are the subject of many current investigations, both in episodic and in semantic memory paradigms (e.g. Brown, 1981).

Within the context of paired-associate learning, we are currently investigating the application of SAM to forgetting paradigms and especially to retroactive and proactive interference studies. This study is being carried out by Ger-Jan Mensink and myself at the University of Nijmegen. In the SAM theory there are two basic reasons why an image may be retrieved better at Time A than at Time B. First, the cues utilized at Time A may be more strongly associated to that image than those used at Time B. Second, the strength and number of other images associated to the cues (even if the cues are the same) may be greater at Time B than at Time A. Everything else being equal, an increase of cue strength to a given image increases both the sampling and recovery probabilities. On the other hand, for fixed cue to image strength, an increase in the strengths of the cue to *other* images will reduce the sampling probabilities (though leaving recovery unaffected).

The increase in the strengths of association of cues to other images tends to be an inevitable consequence of new learning. This new learning will not necessarily lead to forgetting, however. The new information might be organized together or integrated with the old image so strongly that the retrieval of either set of information will then lead at once to retrieval of the other set. This integration could be conceptualized either as resulting in a single, new larger image, or as resulting in two tightly associated images. In the latter case, retrieval of one of the images could result in that information being used as a cue, and thereby eliciting the other image. This possibility is an example of a general principle: forgetting due to new learning occurs when the *same* cue is utilized in an attempt to locate one image among an increasing number of other images. On the other hand, the cues may be changed during the search so that each cue is related to a subset of the increasing number of images; in this event forgetting may be ameliorated or even reversed (leading to proactive or retroactive facilitation).

The decrease in the strengths of association of cues to image can be the result of several factors, chief of which is the change of context over time. The context at the time of storage makes the best retrieval cue, but at the time of test, the context cue used may consist largely of the context information at the time of test, which will usually differ from the storage context by a greater amount as time between storage and test increases.

In order to get predictions for simple forgetting and the effects of retention interval in interference paradigms, we need a model that describes the changes in the strength of contextual associations. For this purpose, we have developed a model for contextual fluctuation that is in the tradition of the fluctuation models developed by Estes (1955) and Bower (1972). We assume that at any given moment the strength of the association between the context cue at test and the stored episodic image is proportional to the overlap in contextual features between cue and image. The basic idea is that context may be represented as a set of contextual features. At any given moment only a small subset of these features is active. This active set of features might be called the 'momentary context'. During retention intervals there is a constant fluctuation between the sets of active and inactive elements. It is assumed that during study the active elements (or part thereof) are encoded in the image. Thus, the amount of overlap will decrease with increasing retention intervals. This leads therefore to a decrease in the strength of the contextual associations. With multiple presentations, new contextual elements are encoded in the image and this accounts (at least partly) for the learning that results from such repeated presentations. One additional complication is that it is assumed that at test the subject will try to reinstate the study context as much as possible. This implies that it will to some extent be possible to direct retrieval to a particular list.

The models that have been (provisionally) developed to deal with interference effects, are quite straightforward given the earlier work. With interfering lists, the retrieval structure contains the images of both lists. Each paired associate is encoded as a single image. Depending on the particular interference paradigm used (e.g. AB-CD or AB-AC), the item cues, which correspond to the stimulus items, are associated strongly to one image (in the AB-CD design) or to two images, one from each list (in the AB-AC design). As mentioned above, separate context cues are used for List 1 and List 2. The List-1 context cue is associated relatively strongly to List-1 images and more weakly to the List-2 images, and similarly for the List-2 context cue. The associative strength between a context cue and an image from the other list, is determined by the overlap between the two presentation contexts. This immediately explains why similarity of presentation contexts leads to more interference. In addition it explains why presenting the lists close together in time, also increases the effect of an interfering list.

The basic interference effects can be predicted quite easily by such a model. Note that in contrast to traditional interference theory no unlearning process is assumed. Retrieval competition due to the sampling process is the basic factor that explains interference in the SAM theory. Because of this, the model has no difficulty explaining the occurrence of proactive interference on socalled MMFR tests, i.e. tests on which the subject is asked to give the responses from both lists. Traditional interference theory cannot explain this result since it assumes that MMFR tests are free of response competition and that proactive interference results from response competition. In SAM this is not the case because the retrieval competition cannot be eliminated simply by instructing the subjects to give both responses. Such a test procedure only eliminates any list discrimination problems the subject might have, i.e. problems in deciding to which list a recovered item belongs. Our retrieval competition explanation is somewhat similar to the list-length effect that occurs in free recall.

Many results from the interference literature can be explained by this model. In this paper, I will mention only a few of them. For the standard interference design the model predicts that there will be a nonmonotonic relation between the number of intrusions from the interfering list and the number of trials on that list, a result that was first obtained in a classic experiment by Melton and Irwin (1940) and has been replicated in many subsequent investigations. In addition, as already mentioned before, the model predicts retroactive interference on MMFR tests. It predicts that as the number of trials on the second list increases, recall from the first list decreases while that from the second list increases. Such results have been interpreted traditionally as demonstrating 'unlearning', the weakening of the first-list associations as a result of the learning of the interfering list. Our analysis shows that such a conclusion is not warranted: these results can be easily explained without using the 'unlearning' concept.

In addition, the model correctly predicts the differences between various retroactive interference designs. It predicts that the AB-CD and the AB-CB designs will be equal to each other and inferior to the control condition. The AB-AC and the AB-ABr designs will also be equal to each other, and both of these will show more interference than the previous two. Precisely such a result has been obtained in a well known experiment by McGovern (1964). It should perhaps be mentioned that these predictions are parameter-free, i.e. they do not depend on the particular parameter values used. In general, we do not want the qualitative predictions to depend on the parameter values unless it can be shown that the result is not always obtained and dependent on the particular details of the experiment. It can be shown that the present model also predicts the independence phenomenon observed by Greeno and others (see Greeno et al., 1978). This independence phenomenon refers to the observation that on MMFR tests in AB-AC designs the probability of recalling the B-response is independent of the probability of recalling the C-response. Even if we introduce reasonable individual differences the predicted correlation is still quite low. Finally, the model predicts a particular relation between the amount of proactive inhibition and the retention interval between the second list and the final test of that list. The predictions show an initial increase in PI followed by a decrease. The increase is due to the fact that as the retention interval increases, the momentary context becomes *relatively* more similar to the List-1 context and this leads to an increase in proactive inhibition. (Note that this also provides an explanation for 'spontaneous recovery', see e.g Postman et al., 1968). The eventual decrease is of course a kind of artefact due to the fact that recall in both the interference and the control condition declines to zero.

In a recent study, Anderson (1981) discovered a particular relation between response latency and response accuracy. It was found that there remains a difference between the control and interference conditions in response latency, even when interference conditions are equated to control conditions in percent recall by additional study trials (see Figure 11). Anderson (1981) showed that such a result is compatible with a particular class of theories. It can be shown quite easily that SAM belongs to this class of theories. In SAM, the probability of recall is a product of the probability of sampling and the probability of recovery. The sampling probability is a function of the *relative* strength to the probe cues, while the probability of recovery is a function of the *absolute* strength. Let R be the relative strength and let A be the absolute strength. We may then write:

$$PC = f_1(R)f_2(A) \qquad [8]$$

where f_1 is the sampling probability and f_2 is the probability of recovery.

If we make the reasonable assumption that the latency of correct responses is determined by the number of sampling attempts made, then it follows that the latency is a function of the relative strength only:

$$RT = f_3(R) \qquad [9]$$

where f_3 is the function that maps the relative strength into reaction time. Since both f_1 and f_3 are monotonic functions of the relative strength R, the following relation holds:

$$f_1(R_1) > f_1(R_2) \quad iff \quad f_3(R_1) < f_3(R_2) \qquad [10]$$

that is, if the sampling probability increases, the latency decreases.

Now let us consider a trial n for the control condition and a trial m for the interference condition such that there is the same probability of recall:

$$f_1(R_{c,n})f_2(A_{c,n}) = f_1(R_{i,m})f_2(A_{i,m}) \qquad [11]$$

Clearly, we have $n < m$. Since the absolute strength is assumed to be proportional to the number of study trials, it follows that $f_2(A_{c,n})$ is lower than $f_2(A_{i,m})$. In order to obtain equality, $f_1(R_{c,n})$ has to be higher than $f_1(R_{i,m})$. In view of the relation between f_1 and f_3, it follows that $f_3(R_{c,n})$ is lower than $f_3(R_{i,m})$. Hence, when the two conditions are equated for percent recall, the response latency will be higher in the interference condition.

This result has an important methodological implication. It shows that according to theories such as SAM, it is impossible to equate the two conditions at the end of second list learning such as is conventionally done by letting the subjects in the two groups learn the second list to an equal criterion. This equalizes the percent recall but does not equalize the absolute stored strengths.

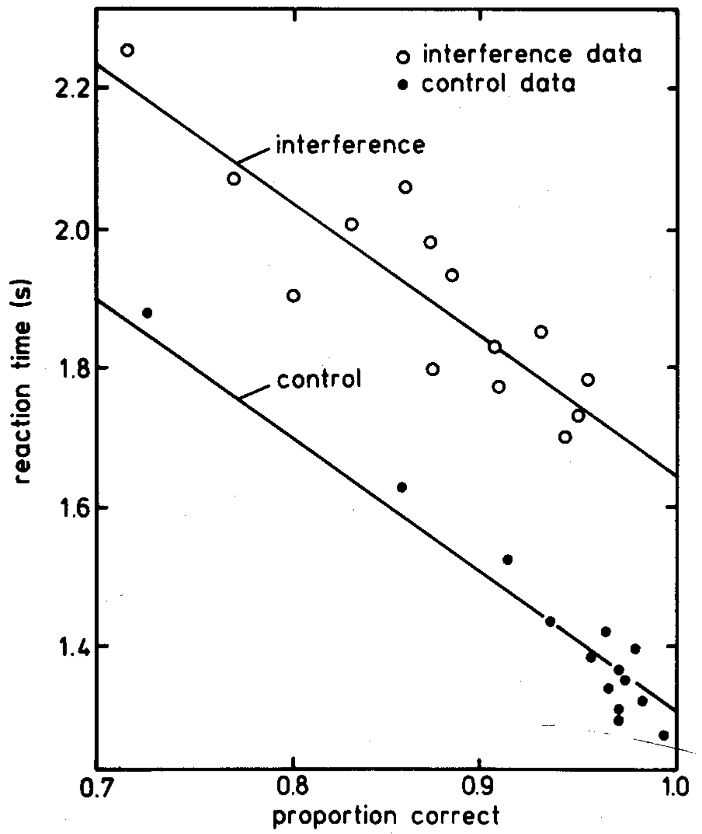

FIGURE 11

Plot of reaction time as a function of proportion correct. (After Anderson, 1981).

4 RECOGNITION

Let us now turn to the application of the SAM model to recognition. Gillund and Shiffrin (1984) have made an extensive analysis of this extension of the model. As mentioned in the introduction, the starting point of this analysis is the assumption that the recognition decision is based on the total activation of LTS in response to the cues. This is equal to the sum of the activations for the individual images, and hence is given by the denominator of the sampling equation:

$$F(Q_1, Q_2, \ldots, Q_m) = \sum_{k=1}^{N} \prod_{j=1}^{m} S(Q_j, I_k)^{W_j} \qquad [12]$$

In simple yes/no recognition tests, it is assumed that two cues are used to probe memory: the context cue, C, and the tested item, W_i. Hence, assuming the weights to be equal to 1, the above equation may be simplified to:

$$F(C, W_i) = \sum_{k=1}^{N} S(C, I_k) S(W_i, I_k) \qquad [13]$$

A positive recognition decision is made when this familiarity value exceeds a criterion value, chosen by the subject. Basically, the storage assumptions are the same as for the free recall model discussed earlier, except that additional variability is introduced to obtain reasonably distributed strength values. The particular variability assumption used has the property that multiplication of all the values in the retrieval structure (before noise is added) by a constant does not change the shapes and overlap of the familiarity distributions for targets and distractors. Because the strengths of a word cue to its own image and to some of the other images on the list will usually be larger than the strengths of a distractor item to those images, the familiarity value wil on the average be higher for list items than for distractors.

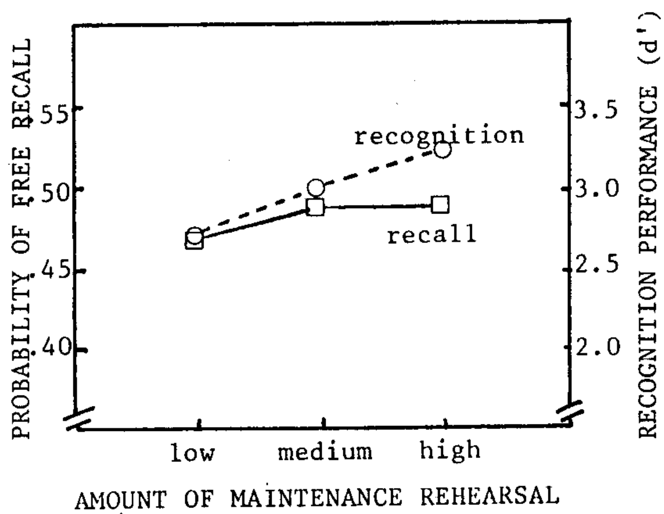

FIGURE 12

Predictions from the SAM model for recognition for increasing amounts of maintenance rehearsal. (After Gillund & Shiffrin, 1984).

Of special interest for any model for recognition is the relation between recognition and recall. Does the model correctly predict which variables have a similar effect on these two retention measures and which variables have a different effect on recognition and recall performance. Gillund and Shiffrin (1984) have documented a number of such predictions of the model. For example, the model correctly predicts a similar effect of list length on recognition and recall. The reason for the decrease in recognition performance with increases in list length, is that increases in list length increase the variance of the familiarity distributions, leading to more overlap of target and distractor distributions. Presentation time also has a similar effect on both measures. In this case, the better recognition performance is due to the increase in the strength of list cues to their own and other list images.

An important aspect of the relation between recognition and recall concerns the differential effects of encoding operations on these measures. For example, it is an established fact in the memory literature that socalled maintenance rehearsal has little or no effect on recall but does have an effect on recognition. It is assumed that maintenance rehearsal causes increases in the amount of self coding, and hence in the strength of an item cue to its own image. In addition, the encoding of contextual features will be slightly increased. No increases are however expected in the interitem strengths. As shown in Figure 12, with these assumptions the model predicts that recall does not change much with increases in the amount of maintenance rehearsal, while recognition does increase. The absence of an effect for recall is predicted because an increase in the strength of an item to its own image increases the tendency of such a cue to sample its own image, thereby reducing the probability to sample an as yet unrecalled item. This negative effect is offset to some extent by the increase in the contextual associative strengths, producing a net effect of little or no decrease. On the other hand, recognition performance is improved by the increases in the strengths of list items to their own images.

A reverse effect is obtained for experimental manipulations that influence the amount of interitem or elaborative rehearsal. Such manipulations have been observed to increase recall performance while having little or no effect on recognition. Figure 13 shows that the present model predicts such results. It is assumed that an increase in the emphasis on interitem coding leads to an increase in the interitem strengths, but to a decrease in the strength of an item to its own image. Both of these factors lead to improved recall performance. In recognition, however, these two factors have opposite effects, and hence this manipulation may lead to a prediction of no net effect.

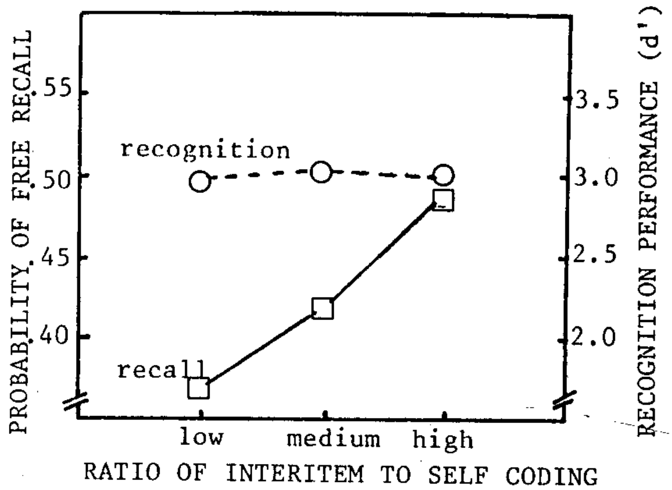

FIGURE 13

Predicted recall and recognition performance as a function of increases in interitem and decreases in self coding. (After Gillund & Shiffrin, 1984).

It has been shown in several experiments that changes in the context between study and test have a negative effect on recall (see e.g. Godden & Baddeley, 1975, 1980; Smith et al., 1978). No effects are however usually found on recognition. Similar results have been reported in the state-dependent literature (see Eich, 1980). These results are predicted by the present model if we assume that context changes decrease the contextual associative strengths. This affects recall but has no effect on recognition (see Figure 14).

A number of other predictions may be derived quite easily from this model. For example, the effects of similarity between list items and distractors can be handled quite easily. Of more interest are the predictions

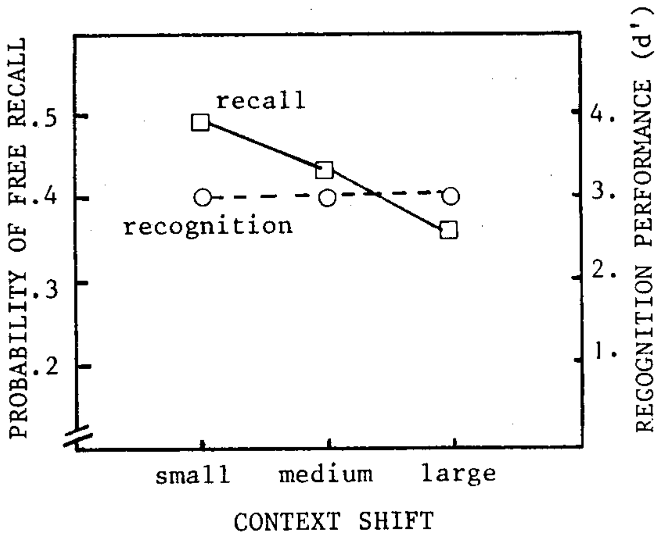

FIGURE 14

Predicted effect of context shifts between presentation and test for recall and recognition. (After Gillund & Shiffrin, 1984).

of the model for the effects of natural-language word frequency on recall and recognition. A classic finding is that high frequency (HF) words are recalled better but recognized worse than low frequency (LF) words. However, this effect is obtained only when pure or unmixed lists are used, that is, when the list consists of either all HF or all LF words. When mixed lists are used, the advantage for HF words is no longer obtained, although they are still recognized worse than LF words.

Gillund and Shiffrin have shown that this interesting pattern of results is predicted by the SAM model. Figure 15 gives the predictions for pure lists. In this application it is assumed that HF cues have higher strengths to the list images than LF cues. Figure 16 gives the results for a mixed list. Why does this model predict no effect of frequency for such a mixed list? In such a situation, we must distinguish between the effects of frequency on cues and to-be-sampled images. Although HF words are better cues than LF words, the probability that a cue samples a HF image is equal to the probability of sampling a LF image. HF and LF words are therefore sampled and recovered equally often, producing equal recall. It may be shown that this SAM model predicts an increase in recall probability as the proportion of HF words on the list increases, although within each such list the probability of recall is equal for both types of words.

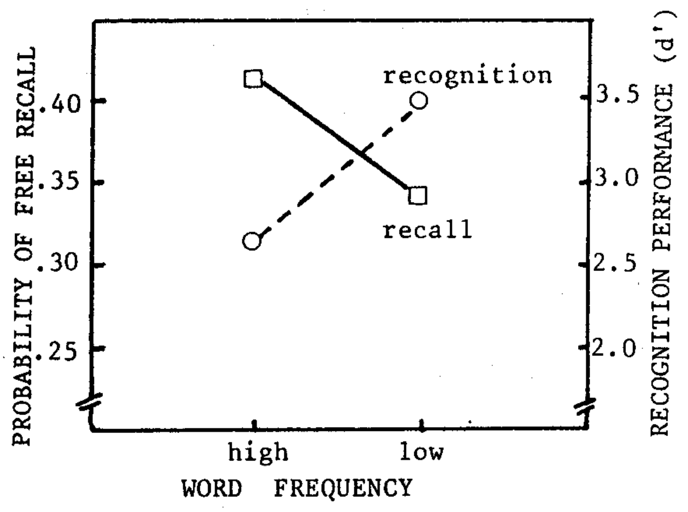

FIGURE 15

Predictions for recall and recognition as a function of word frequency for uniform frequency lists. (After Gillund & Shiffrin, 1984).

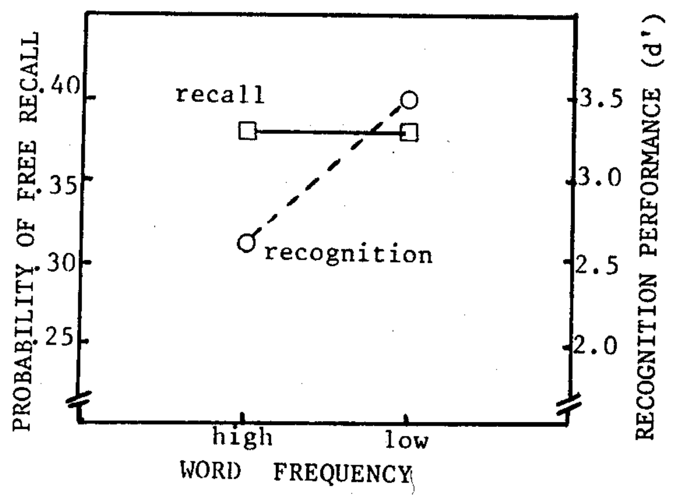

FIGURE 16

Predictions for recall and recognition as a function of word frequency for mixed-frequency lists. (After Gillund & Shiffrin, 1984).

It should perhaps be noted that this model is applicable only to simple recognition tasks. For more complex situations, a more generalized model should be used, in which a recognition decision may be based either on the global familiarity value or on an extended search in which particular images are sampled and recovered. For example, such a generalized model will probably be necessary in those situations in which the list items have been overlearned (see Atkinson & Juola, 1974).

5 FINAL REMARKS

It seems fair to conclude that the SAM theory is able to predict a remarkably large number of empirical phenomena from a variety of experimental paradigms. The SAM theory has been applied to a large number of paradigms, including free recall, cued recall, paired-associate recall, and recognition, as well as to a number of forgetting paradigms. This across-task generality is probably the strongest point in favor of this theory. Because of its generality, SAM is able to integrate results from areas that have traditionally been kept apart.

As a last remark, it might be noted that in all of the applications described above, the predictions from SAM have been used as a test for the correctness of the assumptions on which this framework is based. It is also possible however to use the model as a theoretical framework in which various explanations for a particular phenomenon may be tested against each other. In such applications, one does not test SAM's assumptions but one uses the framework as a tool for theoretical research. Such a use of mathematical models will be of considerable importance given the conspicuous lack of theoretical precision still prevailing in current experimental psychology.

REFERENCES

Allen, M.M. (1969). Cueing and retrieval in free recall. *Journal of Experimental Psychology*, *81*, 29-35.

Anderson, J.R. (1981) Interference: the relationship between response latency and response accuracy. *Journal of Experimental Psychology: Human Learning and Memory*, *7*, 326-343.

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

Basden, D.R. (1973). Cued and uncued free recall of unrelated words following interpolated learning. *Journal of Experimental Psychology*, *98*, 429-431.

Blake, M. & Okada, R. (1973). Intralist cuing following retroactive inhibition of well-learned items. *Journal of Experimental Psychology*, *101*, 386-388.

Bower, G.H. (1972b). Stimulus-sampling theory of encoding variability. In A.W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington, D.C.: Winston.

Bradley, M.M. & Glenberg, A.M. (1983). Strengthening associations: Duration, attention, or relation? *Journal of Verbal Learning and Verbal Behavior*, *22*, 650-660.

Brown, A.S. (1981). Inhibition in cued retrieval. *Journal of Experimental Psychology*, *7*, 204-215.

Eich, J.E. (1980). The cue-dependent nature of state-dependent retrieval. *Memory and Cognition*, *8*, 157-173.

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

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

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

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

Greeno, J.G., James, C.T., DaPolito, F.J. & Polson, P.G. (1978). *Associative learning: A cognitive analysis*. Englewood Cliffs, N.J.: Prentice-Hall.

McGovern, J.B. (1964). Extinction of associations in four transfer paradigms. *Psychological Monographs*, *78*, (16, Whole no. 593).

Melton, A.W. & Irwin, J.McQ. (1940). The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *American Journal of Psychology*, *53*, 173-203.

Mueller, C.W. & Watkins, M.J. (1977). Inhibition from part-set cuing: A cue-overload interpretation. *Journal of Verbal Learning and Verbal Behavior*, *16*, 699-709.

Murdock, B.B.,Jr. (1960). The immediate retention of unrelated words. *Journal of Experimental Psychology*, *60*, 222-234.

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

Postman, L., Stark, K. & Fraser, J. (1968). Temporal changes in interference. *Journal of Verbal Learning and Verbal Behavior*, *7*, 672-694.

Raaijmakers, J.G.W. (1979) *Retrieval from long-term store: A general theory and mathematical models*. Dissertatie KU Nijmegen, Nijmegen.

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). New York: Academic Press.
Raaijmakers, J.G.W. & Shiffrin, R.M. (1981a). Search of associative memory. *Psychological Review*, *88*, 93-134.
Raaijmakers, J.G.W. & Shiffrin, R.M. (1981b). Order effects in recall. In J.B. Long & A.D. Baddeley (Eds.), *Attention and Performance IX*. Hillsdale, N.J.: Erlbaum.
Roberts, W.A. (1972). Free recall of word lists varying in length and rate of presentation: A test of total-time hypotheses. *Journal of Experimental Psychology*, *92*, 365-372.
Roediger, H.L. (1973). Inhibition in recall from cueing with recall targets. *Journal of Verbal Learning and Verbal Behavior*, *12*, 644-657.
Roediger, H.L. (1974). Inhibiting effects of recall. *Memory and Cognition*, *2*, 261-296.
Roediger, H.L. & Schmidt, S.R. (1980). Output interference in the recall of categorized and paired-associate lists. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 91-105.
Roediger, H.L. & Thorpe, L.A. (1978). The role of recall time in producing hypermnesia. *Memory and Cognition*, *6*, 296-305.
Slamecka, N.J. (1968). An examination of trace storage in free recall. *Journal of Experimental Psychology*, *76*, 504-513.
Smith, A.D. (1971). Output interference and organized recall from long-term memory. *Journal of Verbal Learning and Verbal Behavior*, *10*, 400-408.
Smith, S.M., Glenberg, A.M. & Bjork, R.A. (1978). Environmental context and human memory. *Memory and Cognition*, *6*, 342-353.
Tulving, E. (1974). Cue-dependent forgetting. *American Scientist*, *62*, 74-82.
Tulving, E. & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, *5*, 381-391.
Watkins, M.J. (1975). Inhibition in recall with extralist "cues". *Journal of Verbal Learning and Verbal Behavior*, *14*, 294-303.
Watkins, M.J. & Watkins, O.C. (1976). Cue-overload theory and the method of interpolated attributes. *Bulletin of the Psychonomic Society*, *7*, 289-291.