1-1-2005

Modeling the effects of repetitions, study time, and similarity on associative recognition of verbal and nonverbal stimuli

Jing Xu
Iowa State University

Modeling the effects of repetitions, study time, and similarity on associative recognition of verbal and nonverbal stimuli

by

Jing Xu

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Psychology

Program of Study Committee:
Kenneth J. Malmberg III, Major Professor
Veronica J. Dark
Vasant Honavar

Iowa State University

Ames, Iowa

2005

Graduate College
Iowa State University

This is to certify that the master's thesis of

Jing Xu

has met the thesis requirements of Iowa State University

Signatures have been redacted for privacy

TABLE OF CONTENTS

ABSTRACT

To examine the effects of repetition, study time, and similarity on associative recognition (AR), two experiments were conducted and a new model of associative recognition that accounts for both accuracy and latency data was proposed. In Experiment 1, repetition and study time were manipulated within subject and stimulus type was manipulated between subjects. As a package, the results for all three types of stimuli, namely English words, pseudo words, and Chinese characters, disconfirmed single-process familiarity-only or recall-only models and support the predictions of the dual-process model of Malmberg, Holden and Shiffrin (2004). Experiment 2 investigated the latencies of correct AR responses. In contrast to the prediction of the Malmberg et al. model, median latencies for hits were faster than for correct rejections. A new model is proposed to account for both the accuracy and latency of AR.

INTRODUCTION

People discriminate events that occurred from those that did not. In memory research, a classical question is how to characterize this ability, which is called *episodic recognition*. Single-process theories suggest that recognition operates on the basis of familiarity only (e.g., Dennis & Humphries, 2002; Egan, 1958; Gillund & Shiffrin, 1984; McClelland & Chappell, 1998; Murdock, 1982; Shiffrin & Steyvers, 1997). According to these models, when an item is encountered, an episodic memory trace representing the event is stored. At test, the retrieval cue for the stimulus is compared to the contents of memory and a familiarity value is returned, which is compared to a criterion. If the familiarity exceeds the criterion the item is positively endorsed. Because the retrieval cue consists of both context and information corresponding to the item, items that occurred in the given context are more likely to be positively endorsed than items that did not occur in the given context.

Dual-process theories of recognition argue that a recollection process also contributes to recognition decisions (e.g., Atkinson & Juola, 1974; Jacoby, 1991; Malmberg, Holden, & Shiffrin, 2004; Mandler, 1980; Reder et al., 2000; Yonelinas, 1994). When tested with a target, if a previous encounter with the same target in the specified context is recollected, a highly confident positive response will be made; and when tested with a foil, if a previous encounter with a related target is recollected, a highly confident negative response will be made. When recollection fails, a response is based on stimulus familiarity.

There are a variety of recognition tasks that tap different aspects of recognition memory. The simplest recognition task is *single-item recognition*, in which participants study lists of single items, and at test they judge whether a test item was studied. In an

associative recognition task (AR), participants study random pairs of stimuli (e.g. AB, CD, and EF) and usually distinguish intact pairs (e.g. AB) from rearranged pairs (e.g. AD) during the memory test. Thus, AR requires assessing whether two studied items were in fact studied together.

The goal of present study was to investigate the single- or dual-process nature of associative recognition and quantitatively describe it. I begin with a brief literature review of extant theory, especially the experiment and model of Kelly and Wixted (2001), as well as the Retrieving Effectively from Memory (REM) models. Kelly and Wixted's experiments are closely related to the experiments in the current research. After the literature review, I describe Experiment 1, which investigated the effects of repetitions and study time on associative recognition, and the model fitting I have done using the model of Malmberg et al. (2004). I will then describe Experiment 2, which investigated the effects of repetitions and orthographical similarity. The reaction time findings in this experiment have raised a new challenge to the current models of associative recognition. A new REM dual-process model is proposed to account for both latency and accuracy of associative recognition and is fit to the data in Experiment 2.

LITERATURE REVIEW

In this section, I review several key theories of recognition memory, from the global-matching models to the dual-process models. The focus is on the empirical findings and the theoretical accounts of associative recognition. I will then focus on the Retrieving Effectively from Memory (REM) model. The REM dual-process model is the theoretical framework of the current research.

Single-Process Global Matching Models

One prominent class of models is the global-matching models, in which memory decisions are based on a familiarity value computed from the matching of the probe cue against all activated memory traces (e.g. Search of Associative Memory (SAM), Gillund & Shiffrin, 1984; MINERVA 2, Hintzman, 1984; Theory of Distributed Associative Memory (TODAM), Murdock, 1982). Many provide a quantitative description of single-item and associative recognition memory. The SAM and the MINERVA 2 models assume that memory is represented in separate traces. In the SAM model, memory is represented as a matrix of four types of strengths of connection between a probe cue and a memory trace (context strength a, interitem strength b, self-strength c, and residual strength d). Associative recognition is achieved by multiplicative combination of cues, that is, the familiarity value for an intact pair is $F(AB) = 2abc + 2ad^2$ and for a rearranged pair is $F(AD) = 2abd + 2acd$. In the MINERVA 2 model, each item is represented as a vector of random features and each pair is the concatenation of the two item vectors. The familiarity value is calculated by taking the dot product of the probe vector with each vector in memory.

The TODAM model assumes distributive representation of memory. Memory is represented by one composite vector and each studied item is superimposed on the memory

vector. For associative recognition, another vector, the *convolution* of item A and item B, is also added to the memory vector. Thus a memory for a studied pair AB is the weighted sum of each item and the AB association, as shown in equation $\mathbf{M} = \lambda_1\mathbf{a} + \lambda_2\mathbf{b} + \lambda_3(\mathbf{a} * \mathbf{b})$, where **a** and **b** are vectors for item A and B, **a** * **b** is the convolution of item A and B, and λ's are the weights on each vector. A decay factor, the forgetting rate α, is also added to the memory vector, so after studied the *j*th pair, the memory is $\mathbf{M}_j = \alpha\, \mathbf{M}_{j-1} + \lambda_1\mathbf{a}_j + \lambda_2\mathbf{b}_j + \lambda_3(\mathbf{a}_j * \mathbf{b}_j)$. The familiarity value is computed as the dot product of the probe vector and the memory vector, so the familiarity value for an intact pair is $F(AB) = \mathbf{a} * \mathbf{b} \cdot \mathbf{M}_j = \alpha^{j-1}\lambda_3$ and that for a rearranged pair is $F(AD) = \mathbf{a} * \mathbf{d} \cdot \mathbf{M}_j = 0$.

The single-process global-matching models all assume that knowledge of whether the single item was studied (i.e., *item information*) and knowledge of whether the two items were studied together (i.e., *associative information*) are inseparable in memory. However, several findings suggest that item recognition and associative recognition involve functionally distinct mechanisms and/or representations. Gronlund and Ratcliff (1989) found that item information is available about 220 ms sooner than associative information. All the single-process models at that time were challenged by this finding. They suggested a modified single-process model, in which item and associative familiarity values were based on different retrieval cues. *Concurrent cues* consist of individual items and produce familiarity values associated with each item, F(A) and F(B), and a recognition decision can be based on their sum. *Compound cues* jointly match the two items to memory and produce a familiarity value for the AB association, F(AB). Because $F(x)$ is a non-linear function, $F(\mathrm{A}) + F(\mathrm{B}) < F(\mathrm{AB})$. During associative recognition, concurrent information (i.e., item information) is available first because it takes more time to construct the compound cue than the concurrent

cue. Thus only when the compound cue (i.e., associative information) is available are people able to distinguish intact pairs from rearranged pairs.

Hockley (1991, 1992) found that item information is forgotten more rapidly than associative information, which is another challenge to all memory models of associative recognition. He proposed a way to modify the TODAM model in which item information is encoded as content-plus-context features and they sum up to produce a familiarity value for a given item, whereas associative information is encoded as content features of the association only. Context features refer to the physical, spatial/temporal, environmental, physiological, and/or emotional states when the item is presented, whereas content features are item information only, which represent meaning and other high-level codes particular to the studied items (Murnane, Phelps, & Malmberg, 1999). Context changes over time and the difference in the context between study and test is a primary factor in forgetting. This model further assumes that context plays a greater role in single-item recognition than associative recognition, and hence, context changes between study and test result in a faster forgetting for item than associative information.

Recollection in Associative Recognition

Receiver-operating characteristics (ROCs) plot the probability of responding "old" to studied items (hit rate or HR) and new items (false-alarm rate or FAR) as a function of bias to respond "old". Yonelinas (1997) observed that item recognition ROCs are curvilinear whereas associative recognition ROCs are linear. Based on this observation, he proposed a dual-process signal detection model for item recognition and a double high-threshold recall-only model for associative recognition. In the dual-process model for item recognition, recollection and familiarity independently contribute to recognition decisions. Participants

recall a studied item with a probability Ro; the familiarity of a studied item exceeding the decision criterion is a probability Fo, and the probability that the familiarity of a new item exceeding the criterion is Fn. Hence the hit rate is $HR = Ro + Fo - FoRo$ and the false-alarm rate is $FAR = Fn$. The recall-only model of associative recognition assumes that participants recall intact pairs with a probability Ro, and recall that an item on a rearranged pair was not on the same pair during study with a probability Rn. If they do not recall, they guess, and the probability of responding "yes" is G. Thus the hit rate is $HR = Ro + (1 - Ro)G$ and the false-alarm rate is $FAR = (1 - Rn)G$.

Kelly and Wixted (2001; also Dosher, 1984; Gronlund & Ratcliff, 1989) suggested that recollection of associative information and familiarity-based item information both participate in associative recognition. In their experiments, encoding strength was manipulated: Half of the pairs were presented once (weak encoding), and the other half of the pairs were presented six times (strong encoding). Both familiarity-only and recall-only models predict a greater HR in the strong condition than that in the weak condition, but they give opposite predictions on FAR. In most familiarity-based models, the familiarity of intact pairs is a function of the familiarity of individual items plus the familiarity of the association. Thus, increased familiarity with repetitions will increase HRs and FARs. According to the recall-only model, repeating studied pairs would simply increase the probability of the intact pair being retrieved, and hence HRs should increase and FARs should decrease. Kelly and Wixted found a large increase in HRs from weak to strong pairs, but the FARs to weak and strong pairs were nearly equal.

This finding disconfirms most of the familiarity-only and recollection-only models. However, TODAM is an example of a single-process model that predicts no effect of

repetitions on FARs. According to this model AR is based on the familiarity of associative information that is critically assumed to be statistically independent of the items from which the associative information is formed. In TODAM, items are represented by vectors a and b, and the association between **a** and **b** is their convolution **a*b**. If at test a rearranged convolution is used to probe memory, any matches are random.

Based on their findings, Kelly and Wixted (2001) proposed a dual-process model of associative recognition, called the "Some-or-None" model because the associative information is sometimes retrievable, sometimes not. The strength value is either the sum or the difference of item (I) and associative (A) information. When associative information is retrievable, the decision for an intact pair is based on item information plus associative information ($I + A$), and the decision for a rearranged pair is based on associative information subtracted from item information ($I - A$). Otherwise, the decision is based on item information only. Thus the hit rate is $HR = Ro \times [p(I + A) > c] + (1 - Ro) \times [p(I) > c]$ and false-alarm rate is $FAR = Rn \times [p(I - A) > c] + (1 - Rn) \times [p(I) > c]$, where $p(x) > c$ is the probability that information x exceeds criterion, c. Ro is the probability that decision for an intact pair is based on $I + A$, and Rn is the probability that decision for a rearranged pair is based on $I - A$.

The Kelley and Wixted model is a dual-process model in that I information and A information are assumed to be based on different processes. However, it is not a traditional dual-process model in that I information and A information are both continuous random variables. Traditional dual-process models assume that recognition decisions can be based on either a continuous random variable (i.e., familiarity) and/or a discrete random variable (recollection occurs or does not occur).

The REM Model

In this section, I describe the general principles of representation, storage, and retrieval in REM, as well as several versions of the REM model, especially their treatment of associative recognition. The dual-process REM model (Malmberg et al. 2004) is the theoretical framework of the present research.

Representation and Storage

REM models differ in terms of retrieval processes. In general, representation and storage in the different versions of the REM model are the same regardless of the nature of retrieval processes. The REM model assumes that episodic memory consists of separate traces, which include many types of information, such as context, item information, and inter-item information. These traces are represented in REM as vectors of nonnegative integer feature values. Generic knowledge about a studied item is stored in a lexical/semantic memory trace and the occurrence of an event in a certain context is stored in an episodic trace. The features in each vector are represented by non-zero integers drawn from a geometric distribution

$$P(V = j) = (1 - g)^{j-1} g, \qquad j = 1, \ldots, \infty \tag{1}$$

where V is the feature value and g is the parameter for generating the geometric distribution. The geometric distribution allows some feature to be more common than others. A zero value represents no feature is stored.

During the encoding phase, new episodic traces are created by copying the values from the lexical/semantic traces of the studied items. These episodic traces are incomplete and error prone, determined by two parameters: u^* and c. u^* is the probability of storing

any feature and c is the probability of correctly storing a feature value in the episodic trace at a give storage attempt.

When a new event occurrs, if an already stored highly similar trace is retrieved, no new trace will be created and additional features are stored in the same trace; if no similar enough trace is retrieved, a new episodic trace will then be created. This assumption allows the modeling of the repetition and the study-time effects. To keep things simple, it is often assumed that repetition of a studied item will add more features to an existing trace rather than creating a new trace (Shiffrin & Steyvers, 1997, 1998, Malmberg & Shiffrin, in press). The study-time effect is modeled by Malmberg and Shiffrin based on the assumption that storage increases with study time, but the amount of extra storage diminishes over time, as shown in the following equation:

$$t_j = t_{j-1}(1 + e^{-bj}) \tag{2}$$

The equation gives the number of attempts at storing a feature for an item residing continuously in the rehearsal buffer for j seconds, where b is the rate at which storage diminishes with longer study time.

Single-Process Familiarity-Based Retrieval

The simplest version of REM (Shiffrin & Steyvers, 1997, 1998) is a global-matching single-process model. The recognition decision in the REM model is a Bayesian optimal one. It assumes that the lexical/semantic vector of the test item serves as a retrieval cue, which is matched in parallel against all the activated memory traces. For each trace j, a likelihood ratio, λ_j, is calculated as

$$\lambda_j = (1-c)^{n_{jq}} \prod_{i=1}^{\infty} \left[\frac{c + (1-c)g(1-g)^{i-1}}{g(1-g)^{i-1}} \right]^{n_{ijm}}, \tag{3}$$

where n_{jq} is the number of mismatching features in the jth image and n_{ijm} is the number of features in the jth image that match the features in the retrieval cue.

The recognition decision is based on the odds, Φ, the probability that the test item is old divided by the probability the test item is new, which is the mean λ_j:

$$\Phi = \frac{1}{n} \sum_{j=1}^{n} \lambda_j \,. \tag{4}$$

If the odds exceed a criterion, c_t, an "old" response is made; otherwise a "new" response is made. For an optimal decision, $c_t = 1$.

Encoding and Retrieval of Associative Information

The basic principles of REM encoding and representation in associative recognition are the same as described above. More specifically, in several versions of the REM model (Diller, Nobel, & Shiffrin, 2001; Shiffrin & Steyvers, 1997, 1998) pairs are assumed to be the concatenation of the two vectors of features representing each single item in a pair. The accuracy and amount of encoding are still governed by parameters c and u^*, respectively.

Retrieval of pairs in the simplest version of REM is the same as single-process familiarity-based model. The probe cue, the concatenated vector with the first half representing item one and the second half representing item two of the test pair, is matched against all the stored vectored in memory and an odds as the familiarity value is computed according to equations (3) and (4). If the familiarity value exceeds a criterion, an "old" decision is made; otherwise a "new" decision is made.

Using words and faces as stimuli, Criss and Shiffrin (in press) found that in associative recognition, different types of pairs did not interfere with each other during retrieval. This finding can not be predicted by the simplest version of the REM model. To account for this finding, Criss and Shiffrin proposed two ways of representing pairs in REM. In one approach, different pair types are represented by separate regions of the feature vector. For example, there could be six regions: context C, single word W, single face F, word-word pair WW, face-face pair FF, and word-face pair WF. A word-face pair would thus have features stored in the C, W, F, and WF regions. The associative information, such as WW, FF, and WF, is similar to a convolution in the TODAM model. Another approach is called type-code assumption, which assumes that a subset of the feature vector for all pairs is dedicated to encode the pair type. The rest of the vector is simply the concatenation of the two items in the pair. For example, a word-face pair would be represented by WF, the WF-type, and the context features.

The retrieval assumptions in the Criss and Shiffrin's (in press) model still follows the single-process familiarity-based retrieval. In associative recognition, context features are ignored and decisions are based on the odds of associative information, $\Phi_{associative}$.

Dual-Process Retrieval

Malmberg et al. (2004) proposed a dual-process REM model. They suggested that when targets and foils are known to be very similar, a recall mechanism will be invoked after the familiarity value exceeds a certain criterion. Figure 1 illustrates the decision-making steps in this model. A familiarity value is computed in the same manner as in the single-process familiarity-based REM model (Shiffrin & Steyvers, 1997, 1998). If the familiarity value does not exceed a criterion, the response is "new". If it does, the outcome of the

recollection process is examined. If recollection is successful and the recovered details match the stimulus, the response is "old". If the details do not match the stimulus, the response is "new". If recollection fails, recognition decision is based on comparing the familiarity value to another criterion. Notice these assumptions prescribe that recollection-based decisions are in general slower than familiarity-based decisions, because recollection involves more decision steps.

The following equation describes the contribution of recollection to recognition: For an item studied r times, the test item's image is sampled and the feature(s) that discriminates targets from foils is recalled with a probability q

$$\hat{q} = ac\,(1 - (1 - u^*)^{r}), \tag{5}$$

where a determines the recall utilization; c is the probability of correctly storing a feature; u^* is the probability of storing a feature on a given storage attempt; and r is the number of times a target was presented (for similar foils the target item is recalled). The part of this equation in the parenthesis says that the probability of sampling the correct image will increase with r to a level approaching the probability that a feature will be copied correctly.

The dual-process REM model has been motivated by the limitations of the single-process REM model in accounting for the effects of similarity and repetitions in recognition. Figure 2 shows that the single-process REM model predicts a steady rise in FARs with the increase of repetitions and increase in cue similarity. Figure 3 shows that the dual-process model can predict little or no increase in FARs as a function of target presentations (cf. Kelley & Wixted, 2001). In general, the extent to which the recollection process counteracts the familiarity process depends on the tendency to use this strategy, a. At the limit, a is zero, and the dual-process model reduces to the single-process REM model. For relatively high

levels of a, the FAR function tends to bend over in an inverted U-shaped fashion. That is, the discrimination of targets and similar foils improves with increases in the tendency to use a dual-process strategy and in the likelihood of recovering the discriminative feature(s).

The single-process REM model operates on the *differentiation* process to discriminate targets from foils. When an item is repeated, REM assumes that features are more likely to be added to an existing trace until a limit is exceeded. The storage of additional features in an existing trace not only strengthens that trace but it also differentiates it from other traces in memory (McClelland & Chappell, 1998; Shiffrin & Steyvers 1997). That is, the trace becomes less similar to most other traces in memory. The operation of differentiation is based on the assumption that targets and foils are randomly similar. For randomly similar foils, it decreases the FAR and hence improves memory across repetitions because memory traces and dissimilar retrieval cues tend to mismatch more when repetition increases.

However, complexity is added to recognition tasks when the similarity between targets and foils increases. In such situations, the familiarities of both targets and foils are usually greater than the familiarity of dissimilar foils. Similar foils are assumed to share a high proportion, σ, of feature values with the corresponding target, in which there is one list trace that matches a similar foil in almost all its features (Malmberg et al., 2004). In the case of associative recognition, the value of σ is relatively high (Figure 2), because the target and foil pairs share one common word. Increasing target presentations stores more overlapping features of the similar foil and produces greater similar-foil FARs, that is, FAR increases with target presentations. Hence, a single-process strategy is less than optimal in recognition tasks when target-foils similarity is high.

Hintzman, Curran, & Oppy (1992; see also Hintzman & Curran, 1995; Sheffert & Shiffrin, 2003) found that judgments of frequency for targets and highly similar foils (different from targets only in plurality) steadily increase as the number of times a target is studied increases, but discrimination of targets from similar foils did not improve after the first two or three presentations. This phenomenon is known as "*registration without learning*" (Hintzman et al., 1992; Hintzman & Curran, 1995). In such situations as in the registration-without-learning and the associative recognition paradigms when target-foils similarity is high, it may be that matching a critical feature like plurality does not by itself change familiarity enough to allow efficient discrimination of targets from similar foils. Instead, recalling a particular episodic trace might allow the contents of that trace to be examined. Such a recall process counteracts the expected effects of familiarity (see Figure 3; Jacoby, 1991): As repetitions increase, the average trace becomes stronger, causing the familiarity of the similar foil to rise; but recall also improves, which provides an increasingly effective assessment of the critical feature. The result is a correct rejection of the otherwise familiar similar foils.

Based on the above theoretical understanding, the current research further investigates the effects of target-foil similarity and target repetitions in associative recognition. We try to extend the Malmberg et al. (2004) dual-process REM model to these effects in associative recognition. In the following section, I describe the experiments, model fitting, and model development.

EXPERIMENTS AND MODELING

Current Research Approach and Hypotheses

The current project was designed to investigate AR and to extend the REM dual-process model to AR. As described above, manipulating repetitions, study time and target-foil similarity allows one to observe patterns of FARs that are critical for testing AR models. Most single-process familiarity-only models predict a steady increase in FAR with increases in target presentations (but see TODAM), and the recollection-only models predict a decrease in FAR with increases in target repetitions. Dual-process models predict a decrease in FARs to the extent that the recollection process opposes the expected increase in familiarity; hence, FARs may increase, remain steady, or even drop when recollection is particularly effective in counteracting familiarity.

Another objective is to attempt to find a relatively "pure" measurement of familiarity in order to investigate familiarly in isolation from recollection. The approach in Experiment 1 was to use stimuli that are relatively difficult or impossible to recall from a list. Typical memory experiments use word stimuli, but words lend themselves to recollection better than other stimuli because words have well-known labels or because they are meaningful, or both. The difference between meaningfulness and labels can be assessed by comparing performance on words to performance on pseudo words (e.g., *NABLE*) and Chinese characters. For both words and pseudo words, it is possible to generate a phonological label, but pseudo words have no semantic content. Chinese characters, in contrast, have no pronunciation or meaning to those who do not know Chinese.

The hypothesis was that recollection would be less likely to influence the associative recognition of novel nonverbal stimuli. Hence, if we used Chinese characters and pseudo

words to test associative recognition memory, one might find recall to be disrupted and the patterns of HR and FAR would be more like what the familiarity-only models predict, that is, the FAR for rearranged pairs would increase with the number of times the target items were studied. On the other hand, if recollection participates more in the associative recognition of English words, one would find that the FAR decreases, as the single-process recall-only model predicts, or that the FAR remain steady or drop after certain level is reached, as the dual-process model predicts.

An important assumption of the Mamberg et al. (2004) model is that recall is invoked only after the familiarity value exceeds a criterion. As a result foils constructed with randomly similar new items should be quickly rejected based on familiarity information only, whereas the hits to targets should be done after successful recall. As shown in the flow chart of the model (Figure 1), an "old" response will occur either as the result of the recollection process, or after the initial familiarity value fails to give a "new" response, whereas it is possible that some "new" responses are made based on familiarity process only. That means the latency for hits should be greater than that for correct rejections to the "new-new" pairs. In Experiment 2, the target-foil similarity was manipulated on a wider range by introducing three types of "new-new" foils to test this model prediction and observe the similarity effect over a wider range.

Experiment 1

In a pilot study described in my thesis proposal, study time and types of stimulus were manipulated between subjects and repetitions within subject. Pairs of words, pseudo words, or Chinese characters were studied 1, 2, or 6 times, with each presentation lasting for 4 seconds or 1.5 seconds. Participants were asked to discriminate intact pairs from

rearranged pairs at test. A steady level of FARs was expected of across different numbers of repetitions for words. For Chinese characters and/or pseudo words, the data pattern predicted by the single-process familiarity-only model was expected of. That is, FARs were expected to increase with increases in repetitions.

The results showed that the FARs for all three types of stimuli remained relatively constant as a function of repetition in the 4 sec study-time condition. This finding replicated Kelley and Wixted's (2001) findings and extended these results to non-verbal stimuli, albeit at a lower overall level of performance. However, it was hypothesized that Chinese characters would be difficult if not impossible to recollect. The surprising result was that the pattern of FAR for Chinese characters in the 4 sec study condition was similar to words and pseudo words, that is, FAR did not increase with target repetitions. In the 1.5 sec study condition FAR for words also remained invariant across repetitions, whereas FARs for pseudo word and Chinese characters increased with repetitions.

These findings suggest that non-verbal information can be retrieved to correctly reject familiar foils with longer enough study time, and that different stimuli are encoded in different ways (or at different rates), that is, associative information takes longer to encode for Chinese characters than for words or pseudo words. Words are quickly encoded in a manner that facilitates recollection, whereas Chinese characters are encoded more slowly, and hence recollection is only facilitated at relatively long study durations.

An alternative explanation is that study time affected the recognition strategy adopted when Chinese characters were studied. With relatively long study time, participants adopt the dual-process strategy. With relatively short study time, they tend to rely on familiarity-based retrieval strategy. To explore this possibility, study time was varied within subject in

this experiment. The assumption was that participants could not know which pairs were studied for short versus long study durations, and hence a similar retrieval strategy should be used for all test items. Another value of the data obtained from this experiment is that the reduction in the number of between-subjects variable facilitates quantitative modeling, which will be described following this section.

Method

Participants

One hundred and seventy-two non-Chinese speaking students enrolled in undergraduate psychology courses at Iowa State University participated in exchange for extra course credit.

Design and Materials

Participants studied four lists of 24 pairs. Within each list, 8 pairs were studied 1, 2, or 6 times each with spaced repetitions and each presentation of a pair lasted for either 4 sec or 1.5 sec. Pairs of items were constructed randomly and anew for each participant. In each test list, 12 of the tested pairs were intact and 12 of them were rearranged by exchanging the second items in two of the studied pairs.

The type of stimuli comprising the pairs was varied between subjects. Participants studied pairs of English words with a normative frequency of between 20 and 50 per million (Francis & Kucera, 1982), or pairs of pronounceable pseudo words, each constructed by changing one letter in an English word, or pairs of Chinese characters. Forty-two participants studied pairs of English words, 65 participants studied pseudo words, and 65 participants studied pairs of Chinese characters.

Procedure

Participants were tested in individual booths while seated at a desktop computer that controlled the presentation of the stimuli and data collection. Participants received instructions at the beginning of the experiment. They were instructed that they were to study 4 lists of pairs of English words/pseudo words/Chinese characters, and that after each study list a math task was to be performed. A 20-second math task consisted of mentally adding a series of digits. Participants were also informed prior to the experiment that their memory of the studied pairs would be tested after the math task. Participants were informed how the intact and rearranged pairs were constructed and that they were to distinguish the intact pairs from the rearranged pairs by rating on a scale of 1 to 4 how confident they were that the pair was studied. Responses of "1" and "2" respectively indicated high and moderate confidence that the pair was studied, and responses of "4" and "3" respectively indicated high and moderate confidence that the pair was not studied. To make their response, participants were to enter the appropriate number into the computer using the keyboard.

Results and Discussion

The mean HR and FAR for three types of stimuli, two types of study time and three levels of repetitions are listed Table 1 (1.5 sec study time) and Table 2 (4 sec study time). The mean of median Reaction Time (RT) for the correct and incorrect responses are listed in Table 3 (1.5 sec study time) and Table 4 (4 sec study time). Outliers were defined as latencies which deviate from the central tendency by more than 3 SDs for each type of response. The removal of outliers resulted in the loss of 1.8% of the original data points. Figure 4 plots the HR and FAR as a function of number of presentations for the 1.5 sec and 4 sec study-time conditions. An alpha of .05 was adopted as the standard for reliability.

HR and FAR data were first analyzed with an analysis of variance (ANOVA) in which stimulus type (words, pseudo words, and Chinese characters) was a between-subjects variable and study time (4 sec vs. 1.5 sec) and repetition (1, 2, and 6) were within-subject variables. As expected, performance was the best for words and worst for Chinese characters. A main effect of stimulus type was found for HR, $F(2, 169) = 32.005$, $MSE = .079$, $p < .001$. The main effect and linear trend of repetition on HR were significant, suggesting that HR increased with repetitions, $F(2, 338) = 264.740$, $MSE = .028$, $p < .001$, and $F(1, 169) = 431.994$, $MSE = .034$, $p < .001$, respectively. The main effect of study time on HR was also significant, with a higher HR for longer study time, $F(1, 169) = 86.097$, $MSE = .028$, $p < .001$. These main effects were qualified by two interactions: study time × stimulus type, and repetition × stimulus type, $F(2, 169) = 4.243$, $MSE = .028$, $p < .05$, and $F(4, 338) = 2.878$, $MSE = .028$, $p < .05$, respectively. The study time × stimulus type interaction indicated that the add benefit due to extra study time was less for words (5%) than for pseudo words (12%) and Chinese characters (13%). The repetition × stimulus type indicated that the add benefit due to repetitions was more for pseudo words (35% from 1 to 6 presentations) than for words and Chinese characters (28% from 1 to 6 presentations).

For FAR, a main effect of stimulus type was also found, $F(2, 169) = 36.656$, $MSE = .136$, $p < .001$. There were significant main effects of both stimulus type and study time on FAR, $F(1, 169) = 36.656$, $MSE = .136$, $p < .001$, and $F(2, 169) = 5.905$, $MSE = .032$, $p = .016$, respectively. The 3-way interaction among repetition, study time, and stimulus type on FAR was reliable, $F(4, 338) = 5.049$, $MSE = .025$, $p < .01$. Therefore, the FAR data were further analyzed separately for each stimulus type. The main effect of study time was significant for words, $F(1, 41) = 22.974$, $MSE = .016$, $p < .001$, but not for pseudo words and

Chinese characters, $F(1, 64) = .514$, $MSE = .040$, $p = .476$ and $F(1, 64) = 1.333$, $MSE = .035$, $p = .252$, respectively. The main effect of repetition was marginal for words, decreasing by 6% from 1 to 6 target presentations, $F(2, 82) = 2.645$, $MSE = .030$, $p = .077$; for pseudo words, the main effect of repetition was significant and in the opposite direction, increasing by 8% from 1 to 6 target presentations, $F(2, 128) = 5.775$, $MSE = .038$, $p < .005$; and for Chinese characters, the main effect of repetition was also marginal, increasing by 5% from 1 to 6 target presentations, $F(2, 128) = 2.388$, $MSE = .040$, $p = .096$.

In the 4 sec study-time condition, ANOVAs with repetition as the within-subject variable for all three types of stimulus revealed that repetition did not reliably affect FAR, $F(2, 82) = 1.779$, $MSE = .017$, $p = .175$ for words, $F(2, 128) = 2.552$, $MSE = .030$, $p = .082$ for pseudo words, and $F(2, 128) = .371$, $MSE = .034$, $p = .691$ for characters. In the 1.5 sec study-time condition, repetition did not reliably affect FAR for words, $F(2, 82) = 2.548$, $MSE = .034$, $p = .084$. However, repetition did reliably affect FARs for pseudo words and Chinese characters in the 1.5sec study-time condition, $F(2, 128) = 12.208$, $MSE = .033$, $p < .001$ and $F(2, 128) = 7.745$, $MSE = .034$, $p < .005$, respectively.

Further analysis showed that the data pattern predicted by the dual-process model was obtained for the pseudo words and Chinese characters groups, that is, FAR increased as repetition increased until a certain point was reached. The pattern is described more fully in the following section. A linear trend of repetition on FAR was found in the 1.5 sec study-time condition for both pseudo words and Chinese characters, $F(1, 64) = 21.392$, $MSE = .034$ $p < .001$, and $F(1, 64) = 13.983$, $MSE = .034$, $p < .001$, respectively. FAR for pseudo words appeared to continue increasing to at least 6 target presentations, because t-tests showed that the FARs for pseudo words significantly increase from 2 to 6 target presentations $t(64) = -$

3.500, $SE = .033$, $p = .001$. However FAR for Chinese characters appeared to increase from 1 to 2 target presentations, $t(64) = -2.981$, $SE = .032$, $p = .004$, and then leveled off from 2 to 6 target presentations, $t(64) = -.774$, $SE = .033$, $p = .442$.

Another noteworthy finding is that participants appeared to have adopted a more lenient criterion to pseudo words than to words. The criterion location c_a was calculated using the equation suggested by Macmillan and Creelman (2004). The mean c_a for words ($M = -.009$, $SE = .036$) was higher than that for Chinese characters ($M = -.085$, $SE = .032$), and than that for pseudo words ($M = -.138$, $SE = .040$). T-tests detected significant difference between the c_a's for words and pseudo words but no significant difference between the c_a's for words and Chinese characters, $t(104) = 2.218$, $SE = .058$, $p = .029$, and $t(103) = 2.218$, $SE = .058$, $p = .131$, respectively.

In summary, Experiment 1 replicates the findings in the pilot study. AR for words was better than for pseudo words and AR for pseudo words was better than for Chinese characters. HR for three types of stimulus increased with both study time and repetitions. FARs for three types of stimulus stayed relative constant in 4 sec study-time condition. However, with short study time (1.5 sec) only FAR for words was unaffected by repetitions, whereas FARs for pseudo word and Chinese characters increased with increases of repetitions. A detailed interpretation of these findings is given in the following section of model fits using the Malmberg et al. (2004) dual process REM model.

Model Analysis of Experiment 1

Experiment 1 found that participants' performance was the best for words and the worst for Chinese characters. One explanation of the performance difference among three types of stimulus within the present theoretical framework is that words are encoded more

effectively than pseudo words and pseudo words are encoded more effectively than Chinese characters. For English speaking participants, words have inherent perceptual, phonological, and semantic information; pseudo words do not have semantic information, whereas Chinese characters have only perceptual information. Thus, more information is encoded in a given unit of time when a word is studied than when a pseudo word is studied, and more information is encoded in a given unit of time when a pseudo word is studied than when a Chinese character is studied. There are a number of ways to model this assumption. Based on Equation 5, it is likely that study time affects the encoding of each stimulus type differently. As encoding is reduced (i.e., u^* or c approaches 0), recollection, q, is reduced, and it becomes more difficult to discriminate intact and rearranged pairs. The simplest assumption is that the amount (u^*), instead of the accuracy (c), of information encoded differs for the three types of stimuli. Following this assumption the u^* value for words was assigned the value of 0.06, for pseudo words the value was 0.03, and for Chinese characters the value was 0.013.

Another finding in Experiment 1 is that study time significantly affects the performance of AR for three types of stimulus. Both HRs and FARs for three types of stimulus were higher in the short study-time condition than in the long study-time condition. This study-time effect was modeled by the assumption that storage increase with study time, but the amount of extra storage diminishes over time, as shown in Equation 2. For simplicity, the diminishing factor, b, was fixed at 1.0, and the number of storage attempts for one second presentation t_1 was set at the value of 7.

Then for the repetition effect, the simple assumption in the current model was that spaced repetitions produce a single episodic image and that the total number of storage

attempts is the product of the number of repetitions and the number of attempts in one presentation. The repetition effect found in Experiment 1 is most clearly demonstrated in the data pattern of HRs. HRs for three types of stimulus increased with increases of repetitions in both 1.5 sec and 4 sec study-time conditions.

Perhaps most interesting aspect of the present model is the assumption that a recall-to-reject strategy was used for pseudo words and Chinese characters. The flat pattern of FARs for nonverbal stimuli is similar to those found for words in the 4 sec study-time condition, suggesting that the recall-to-reject strategy can be utilized for a variety of stimuli. Parameter *a* in Equation 5 represents the tendency to use a recollection strategy. The recall-to-reject strategy was assumed to be more emphasized for verbal stimuli than non-verbal stimuli. Thus, in the model, for words, a = 1.0, and for pseudo word and Chinese characters, a = 0.8.

Lastly, Greene (2004) reported that people were more likely to respond "old" to non-verbal stimuli than to words. The current results extend these findings to AR and to Chinese Characters. Bias is modeled in REM by using a different criterion for words vs. nonverbal stimuli. Figures 4 c. and 4 d. show that the REM model qualitatively and quantitatively captures the data from Experiment 1.

In summary, the data patterns for HRs and FARs found in Experiment 1 can be summarized into three effects: the stimulus effect, the study time effect, and the repetition effect. To be specific, it was found that HRs for three types of stimulus increase with increases of target repetition and study time; FARs were the lowest for words and the highest for Chinese characters; and FARs for words stayed constant in both 1.5 sec and 4 sec study-time conditions and FARs for pseudo words and Chinese characters also stayed invariant in

the 4 sec study-time condition, but they continues increasing from 1 to 6 and from 1 to 2 target presentations respectively in the 1.5 sec study-time condition. These effects were modeled based on the assumptions of that encoding is the most effective for words and the least effective for Chinese characters, that retrieval strategy is more emphasized for words than non-verbal stimuli, and that participants adopted more stringent criteria for words than non-verbal stimuli.

Experiment 2

In a typical associative recognition paradigm, similarity between targets and foils is usually assumed to be relatively high, because the target and the rearranged foil share one common word. For example, if two targets in the study list are, "*ARMS YEAR*", and "*BREAD NORTH*", the rearranged pairs in the test list would be "*ARMS NORTH*", and "*BREAD YEAR*." In Experiment 2, target-foil similarity was manipulated over a wider range by manipulating the perceptual similarity of the targets and foils. This manipulation was achieved by using A'B' and A'D' foils, as well as the XY foils. The A'B' and A'D' pairs were orthographically similar to targets and rearranged pairs, respectively. For example, an A'B' pair generated from the target "*ARMS YEAR*" could be "*ARTS BEAR*" and an A'D' pair generated from targets "*ARMS YEAR*", and "*BREAD NORTH*" could be "*ARTS WORTH*", The XY pairs were randomly similar to targets. The results from Experiment 1 suggested that recall-to-reject strategy was also used in pseudo words and Chinese characters, that is, lower-level nonverbal information might be used in the recollection process during AR, especially as the mechanism for rejecting otherwise familiar pairs of pseudo words or Chinese characters. In Experiment 2, the difference found in FARs for different types of foil will provide an estimate of the use of higher-level semantic codes as means for sampling

and recovering episodic information used to reject otherwise familiar foils, because the targets and foils, especially AB and A'B' pairs, are perceptually so similar that people may have to rely more on their difference at the semantic level to discriminate them apart.

Further, the combination of the patterns of FARs and the latencies of correct responses also will constrain the possible models. For instance, while the dual-process model of Malmberg et al. (2004) can adequately handle the accuracy data from Experiment 1, it is unknown whether it is adequate for explaining the latency data in the AR task. According to that model the recollection mechanism is invoked after familiarity exceeds a certain criterion. Thus, one should expect that it would typically take more time for an "old" response to targets than a "new" response to a randomly similar foil. That is, the average hit should be slower than the average correct rejection of items only randomly similar to targets (XY pairs).

Method

Participants

Forty-two students enrolled in undergraduate psychology courses at Iowa State University participated in exchange for extra course credit.

Design and Materials

As in Experiment 1, an associative recognition experiment paradigm was used, in which participants were asked to discriminate foil pairs from target pairs. There were four types of foil pairs: 1) rearranged pairs that were constructed from the study list by exchanging the second items in two of the studied pairs (AB, CD → AD, CB); 2) new pairs that were orthographically similar to the target pairs (AB → A'B'); 3) new pairs that were orthographically similar to the rearranged pairs (AB, CD → A'D', C'B'); and 4) new pairs

that were randomly similar to the target pairs (XY). English words that were orthographically similar (e.g. bike, bite) were selected from the list used by Shiffrin et al. (1995) and each pair of similar words shared a vowel sound and exactly one consonant cluster.

Participants studied six lists of 18 pairs each. The presentation time of each pair was 4 seconds and six pairs in each list were studied 1, 2, or 6 times with spaced repetitions. Each test list was composed of 12 targets (intact pairs) and 12 foils with 6 XY pairs and either 6 A'B' or A'D' pairs.

Procedure

Participants were tested in individual booths while seated at a desktop computer that controlled the presentation of the stimuli and data collection. Participants received instructions at the beginning of the experiment. They were instructed that they were to study six lists of pairs of words with each pair appearing on the screen for four seconds, and after each study list a math task was to be performed. A 20-second math task consisted of mentally adding a series of digits. Participants were also informed prior to the experiment that their memory of the studied pairs would be tested after the math task. The tested pairs were intact, rearranged, or new pairs. Participants were informed how the intact and rearranged pairs were constructed and that they were to distinguish the intact pairs from the rearranged and the new pairs by rating on a scale of 1 to 4 how confident they were that the pair was studied. Responses of "1" and "2" respectively indicated high and moderate confidence that the pair was studied, and responses of "4" and "3" respectively indicated high and moderate confidence that the pair was not studied. To make their response, participants were to enter the appropriate number into the computer using the keyboard.

Results and Discussion

Accuracy data

The mean HR and FAR for the five pair types and three levels of strengths (1, 2, and 6 presentations) are listed in Table 5. Figure 5a plots the HRs and FARs as a function of presentations. An alpha of .05 was adopted as the standard for reliability.

HR data were analyzed with an ANOVA in which number of presentation (1, 2, and 6) was a within-subject variable. Similar to Experiment 1, there was a reliable effect of number of presentations, $F(2, 82) = 68.494$, $MSE = .006$, $p < .001$. The mean HR appeared to increase when the number of presentation increased.

FAR data were first analyzed with an ANOVA with two within-subject variables, repetition (1, 2, and 6) and foil type (AD, A'B', and A'D'). There was a main effect of repetition on FAR, $F(2, 82) = 29.815$, $MSE = .044$, $p < .001$. However, the main effect of foil type on FAR was only marginal, $F(2, 82) = 2.590$, $MSE = .037$, $p = .081$. Thus, the orthographic similarity effect on FARs was not as strong as expected. A one-way ANOVA on the FAR was done for each type of foil including the XY pairs with number of presentation (0, 1, 2, and 6) as a within-subject variable to see if any of the foils "similar" to target (AD, A'B', and A'D') were treated differently from the brand "new-new" pairs. A'D' and XY pairs were not significantly different from each other, $F(3, 123) = 1.314$, $MSE = .016$, $p = .273$. The same result was found with the FARs for A'B' and XY pairs, $F(3, 123) = .877$, $MSE = .014$, $p = .455$. However, there was a difference between the FARs for AD and XY pairs, $F(3, 123) = 8.014$, $MSE = .05$, $p < .001$. T-test showed that even with 1 presentation FAR to AD pairs was significantly greater than that for XY pairs, $t(41) = 4.057$, $SE = .04$, $p < .001$. T-tests also showed that FAR for XY pairs was not significantly different

from A'B' and A'D' pairs, $t(41) = .864$, $SE = .02$, $p = .393$ and $t(41) = .507$, $SE = .02$, $p = .615$, respectively. These results appear to suggest that all the "new-new" pairs (A'B', A'D', and XY pairs) were treated as the same type of foil pairs. One possibility is that only semantic information is used to perform AR, but the results of Experiment 1 suggest otherwise, and as will be seen below, the analysis of the latency data is informative.

Another interesting finding was that although, similar to the findings in Experiment 1, FAR for rearranged (AD) pairs stayed invariant as a function of repetition, $F(2, 82) = 1.335$, $p = .269$, the mean FAR for AD pairs ($M = .292$, $SE = .037$) is much higher than that for words under 4 sec study-time condition in Experiment 1 ($M = .160$, $SE = .019$). An explanation of that could be that the inclusion of new-new pairs, which consisted of the 7/8 of the total foils pairs, affected peoples' retrieval strategy. They may have tended to rely more on the familiarity evidence to make their decisions and thus gave more old responses than in Experiment 1. This effect becomes more obvious when the reaction time (RT) data are examined.

Latency data

The means of median Reaction Times (RTs) for the correct and incorrect responses are listed in Table 6. Outliers were defined as latencies which deviate from the central tendency by more than 3 SDs for each type of response. The removal of outliers resulted in the loss of 1.5% of the original data points. Figure 5b plots RTs for correct responses as a function of presentations. An alpha of .05 was adopted as the standard for reliability.

Repetition effect was found on hit RTs, but not on correct-rejection (CR) RTs. An ANOVA analysis done on the hit RTs with repetition as a within-subject variable detected a significant linear effect of repetitions, that is, the hit RTs decreases with the increase of

repetitions, $F(1, 41) = 44.235$, $MSE = .087$, $p < .001$. An ANOVA on RT data for CRs with repetition (1, 2, and 6) and foil type (AD, A'B', and A'D') as two within-subject variables did not detect a main effect of repetition, $F(2, 70) = .556$, $MSE = .670$, $p = .576$.

The similarity effect was found on the RT data of the CRs to four types of foil pairs. An ANOVA on the RT data for CRs with two within-subject variables, repetition and foil type, detected a main effect of foil type, $F(2, 70) = 41.457$, $MSE = .801$, $p < .001$. Participants were the slowest in making CRs to an AD pair ($M = 3.077$, $SE = .161$), which shared the highest similarity with the targets; participants were faster in making CRs to an A'B' pair ($M = 2.224$, $SE = .132$), which was a new-new pair but both words shared some orthographical similarity with a target pair; they were even faster to an A'D' pair ($M = 2.036$, $SE = .108$), which was also a new-new pair and only one word was orthographically similar to a word on a target pair. A t-test on the CR RTs to XY pairs and A'D' pairs revealed that CR RTs to XY pairs are not significantly different from those to A'D' only when the corresponding targets were repeated once, $t(41) = .924$, $SE = .08$, $p = .361$. That may suggest that the similarity effect takes on after two repetitions of the target.

The most interesting finding in this experiment is that a t-test on the CR RTs to XY pairs and the hit RTs showed that RTs for the CRs to XY pairs was at the same level as those of hits to the target pairs presented only once, $t(41) = .742$, $SE = .07$, $p = .463$. After two presentations, hit RTs were faster than the CR RTs to new pairs, $t(41) = -3.918$, $SE = .04$, $p < .001$ and $t(41) = -6.150$, $SE = .06$, $p < .001$ for 2 and 6 presentations respectively. This finding can not be predicted by Malmberg et al.'s (2004) model, since it assumes that hits would happen only after recollection was attempted, whereas the CRs to the XY pairs are usually quickly rejected based on their familiarity.

In summary, Experiment 2 found that HR increases whereas RTs for hit decrease with the increase of repetitions; but repetition did not reliably affect either FARs or RTs for CRs to four types of foils. Orthographic similarity manipulation did not affect FARs to A'D', A'B' and XY pairs but did affect RTs for CRs to four different types of foils. CRs for AD pairs were the slowest, whereas CRs for A'D' and XY pairs were the fastest. The most significant finding was that hit RTs became faster than CR RTs for XY pairs after targets being presented for more than once. This finding disconfirms the dual-process REM model (Malmberg et al., 2004), which assumes that the slower recollection process is invoked only after the familiarity value exceeds a criterion when the similarity between targets and foils is high. A detailed interpretation of these findings is given in the section of Model Fits of Experiment 2, using a new model that accounts both accuracy and latency data of associative recognition.

A Model for both Accuracy and Latency in Associative Recognition

To account for both the accuracy and the latency data in Experiment 2, a new model was developed that assumes that the familiarity and recollection processes run in parallel and recognition decisions are made based on either familiarity or recollection evidence. In order to account for the latency of AR the time course of the retrieval processes and how decisions are made is described. Before I describe the new model, I will give a brief review of the key models that account for both speed and accuracy performance in recognition memory.

An Overview of Speed and Accuracy Models of Recognition Memory

The Atkinson & Juola model. Atkinson and Juola (1974) proposed the first dual-process recognition memory model to account the relationship between speed and accuracy in performance. In this model, the recognition decision is assumed to be based on both

familiarity and recollection. The familiarity information is a continuous random variable with a mean value higher for the old items than for the new items. There are two decision criteria along the familiarity continuum. If the initial familiarity value for a test item is above the higher criterion, a fast positive response is given; if it is below the lower criterion, a fast negative response is given. If it falls in between the two criteria, an extended memory search, that is, recollection, is executed before the recognition decision is made. The extended search will guarantee that a correct response will be made, and the time it takes is in proportion with the length of the study list. Thus, the speed-accuracy trade-off phenomenon can be captured by adjusting the distance between the two decision criteria. The larger the distance between the two decision criterion, the higher the performance and the longer the mean reaction time.

Random walk models. The random walk models constitute a large branch of the *sequential sampling models*. The sequential sampling process is the accumulation of evidence from successive samples of noisy stimulus representation over time until a criterion amount of information is exceeded (Ratcliff & Smith, 2004). Two important features of the sequential sampling process are that accumulation of evidence occurs over time and that the process is stochastic. The random walk models assume that evidence for and against the alternative choices accumulate over a sequence of time steps and compete against each other. Once the evidence for one of the choices exceeds that for another by a criterion amount, a decision is made.

Ratcliff (1978) proposed the *Wiener Diffusion model* to account for both accuracy and reaction time in recognition memory. It assumes that information accumulates with a certain rate over time, moving the evidence level up and down toward one of the two response

boundaries. The accumulation of evidence is achieved through the *diffusion* process, in which accumulation happens in infinitely small steps over continuous time. Once the evidence level reaches the upper boundary, a positive response is given; if it reaches the lower boundary, a negative response is given. The rate that information accumulates can be positive or negative, depending on the difference between a criterion and the amount of feature overlap between the test stimulus and the memory storage. The interaction of the characteristics of information accumulation and the distance between the two boundaries captures the relationship between speed and accuracy of performance: with the same boundary locations, RTs will be longer and accuracy lower when accumulation rate is slower or less amount of information is accumulated at each step; with the same evidence accumulating characteristics, the closer the two boundaries, the faster but less accurate the responses are.

Another model that belongs to the random walk class is the *Ornstein Uhlenbeck (OU) model* (Busemeyer & Townsend, 1992, 1993). In addition to the assumptions of the diffusion model, a decay mechanism is added in the OU model: the rate of evidence accumulation is counteracted by a decay force, which is a function of the amount of evidence accumulated. The decay mechanism has been introduced to account for the asymptotic accuracy in response signal experiments (Usher & McClelland, 2001).

Accumulator models. The accumulator models belong to another large branch of the sequential sampling models. This class of models assumes that information accumulates toward both alternative choices in parallel, each being monitored by a counter, which could be mutually independent or correlated to a certain degree. A response is made when one of the counters passes an absolute criterion. One of the accumulator models is the *Poisson*

counter model (Pike, 1973; Townsend & Ashby, 1983). In this model, the two counters for positive and negative evidence are independent from each other. Evidence arrives at each counter in units at exponentially distributed time intervals with two different rates. Therefore the evidence streams are Poisson processes. The difference of the two rates represents the quality of stimulus information, so different types of stimuli have different drift rates. When one of the counters reaches its threshold, a response is given according to which counter it is. Van Zandt, Colonius, and Proctor (2000) have shown that the Poisson counter model can successfully account for the RT data in perceptual matching tasks.

Another important accumulator model is the *leaky, competing accumulator model* (Usher & McClelland, 2001). In addition to the assumption of stochastic information accumulation over time in sequential sampling models, this model also assumes that the accumulation process is subject to leakage, or decay and that evidence toward different choices compete against each other through a lateral inhibition process. As evidence is added to one accumulator, other accumulators are inhibited. Once the evidence in one accumulator passes its absolute criterion, a response is generated. In this way, the model extends naturally from two-alternative choices to multiple choices. Another two assumptions of this model are recurrent self-excitation and nonlinearity of information accumulation. The architecture of the model takes the form of a neural network with two layers: the input units and the accumulator units. The activation of an accumulator, x_i, is captured in the following equation:

$$\begin{aligned} dx_i &= [\rho_i - kx_i - \beta \textstyle\sum_{i' \neq i} x_i] \frac{dt}{\tau} + \xi_i \sqrt{\frac{dt}{\tau}} \\ x_i &\rightarrow \max(x_i, o) \end{aligned} \tag{6}$$

where τ is a time scale, $\xi_i$is a Gaussian noise term for the decay factor, ρ_i is the external input, k is the scaling factor for recurrent self-excitation, and β is the inhibition weight.

The Ballistic Leaky Competitive Accumulator (BA) model. Brown and Heathcote (2005) recently proposed a ballistic accumulation model, which is the first ballistic model that can successfully account for RT data from choice tasks. The BA model is the simplified leaky competitive accumulator model (Usher & McClelland, 2001) with the stochastic component dropped. The stochastic models assume that accumulated evidence varies at each time step during one decision trial. In the BA model, the accumulation process during each trial is deterministic, that is, instead of being a sequential sampling process, it is a ballistic process. The same as the leaky competitive accumulator model, the BA model also assumes the trial-to-trial variability in input strength and starting point of evidence accumulation.

The Memory Interrogation model (MIM). Hockley and Murdock (1987, 1992) proposed the MIM decision model to account for the accuracy and latency data as an alternative to the sequential sampling models. In this model, information does not accumulate. Instead, it is the decision process that takes time. The input of the decision process is the value of match between a test probe and memory, which is the result of the same matching process as in the TODAM model. The decision process is carried out with cycles. As in most latency models, there are two decision criteria in this model. During each cycle, a certain amount of noise is added to the matching value. A decision is made if the new matching value exceeds the upper criterion or falls below the lower criterion. If the value falls in between the two boundaries, a new decision cycle begins. In this new cycle, the two criteria are moved closer to each other by a constant and a new sample of noise is added to the original matching value. The decision cycles continue on until one of the

decision criteria is reached. It is also assumed that each successive cycle takes longer and longer time, as shown in the following equation:

$$Decision\ Latency = (k^2 + k + 2)BCT \tag{7}$$

where k is the number of cycles and BCT is the base cycle time. The response latency is the decision latency plus the time for other processes, drawn from a normal distribution.

The Current Model Assumptions

A new REM dual process model for associative recognition is proposed in an attempt to account for both the accuracy and latency data found in Experiment 2. It combines the traditional SAM-REM (Gillund & Shiffrin, 1984; Shiffrin & Steyvers, 1997) models with the diffusion model by assuming the following:

1. Encoding is modeled in the same way as in Experiment 1, because memory is represented in the same way in all versions of REM model. The encoding accuracy parameter c is set at the typical value 0.7, and the encoding amount parameter u^* is set at 0.04. Again, to model the repetition effect, it is assumed that spaced repetitions produce a single episodic image and that the total number of storage attempts is the product of the number of repetitions and the number of attempts in one presentation. The number of attempts for one second presentation t_1 is 7.
2. The retrieval process occurs over time as successive cycles. For each unit of retrieval time, t, a set of features, F_t, is sampled with replacement and a probability of s from an activated memory set.
3. Retrieval is based on both familiarity and recollection processes. The familiarity-based process operates in the same fashion as the single-process REM model. For

each t, global matching is performed over F_t producing the familiarity value Φ_t. To produce both positive and negative evidence that the test pair was studied as in most latency models, the logarithm of Φ_t is computed. At a certain time, t_m, a familiarity based recognition decision is based on accumulated values of odds up to time m:

$$\Phi_m = \frac{1}{N}\sum_{t=1}^{m} \log \Phi_t \,. \qquad (8)$$

4. The recall process involves the same mechanisms of sampling and recovering traces as in the SAM model (Malmberg & Shiffrin, 2005). The sampling process follows the Luce choice rule. The probability of sampling trace, i, from the set of sampled features, F_t, given retrieval cue, Q, is positively related to the similarity of Q and trace i, and negatively related to the similarity of Q and other traces in memory:

$$P(i \mid Q) = \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j} \,. \qquad (9)$$

The recovery is a threshold process. Recovery of trace i is successful only when there are at least k correctly stored features given F_t.

5. The recollection process is slower than the familiarity-based process. Familiarity evidence accumulates over all the units of t, but sampling and recovery occurs less frequently. That is, familiarity continues to accumulate while the slower sampling and recovery processes are taking place. For simplicity, it is assumed here in the current modeling that sampling occurs every other t beginning at $t = 2$.

6. A recognition decision is made only when evidences for both familiarity and recollection are available. An "old" decision is made if an AB (target) pair is tested and the AB trace is recalled or if $\log \Phi_m$ exceeds a positive "old" criterion. A "new" decision is made if an AD (rearranged) pair is tested and the AB or CD is recalled or if Φ_m less than a negative "new" criterion. Recognition decisions are made for those time units, *t*s, on which a trace has been sampled.

Model Analysis of Experiment 2

The new REM model's prediction to the accuracy and latency data in Experiment 2 is plotted in Figure 5c and 5d, respectively. The model gives describes the accuracy data pretty well. The critical finding in this experiment is that the latency for the correct rejections is almost the same as the latency for hits to targets presented once. The model gives the same prediction. The model also predicts that the latency of hits will decrease with more repetitions, the same as shown in the data.

The model also predicts that the correct rejection RT to AD pairs decreases with increased number of target presentation, although the data shows that these RTs are not significantly different from each other. This is a part of the model that demands modification.

The similarity effect is modeled by manipulating a new parameter for orthographical similarity, which is operationally defined as the proportion of overlapping or common features between a target and a foil. Here, the value is set at 0.3 for orthographically similar pairs because it leads to reasonable fits, and since XY pairs are randomly similar to targets it is set to 0. This manipulation results in the highest similarity between rearranged pairs (AD) and targets (AB), and the lowest similarity between brand "new-new" pairs (XY) and targets

(AB). The orthographically similar "new-new" pairs are in between, that is, they are less similar to targets than AD pairs, but more similar to targets than XY pairs. A'B' pairs are more similar to targets than A'D' pairs are due to the nonlinear activation function (Eq. 3).

The old and new criteria and the sample size are the three most interesting parameters in the new REM model. As the distance between the two criteria increases, recollection participates more in recognition and it takes longer time before a decision is made. In the current simulation, the old and new criteria are set at 1.0 and -0.4 respectively. That is, more familiarity evidence is needed for an old decision than a new decision. The parameter for sample size, the probability s for sampling a set of features F_t at each time cycle, also affects the rate of evidence accumulation and the extent that recollection participates in recognition. As the size increases, more recollection happens and evidence accumulates faster. In the current modeling, sample size is set at 0.26. That is, at each unit of time, with a probability of 0.26 that the features stored in memory will be sampled.

GENERAL CONCLUSIONS

In the present research, we did two experiments to examine the effects of repetition, study time, and similarity on associative recognition. The findings in these two experiments as a package are difficult for most of current models of associative recognition to handle. Although the dual-process REM model of Malmberg et al. (2004) has been successfully extended to Experiment 1, the RT data from Experiment 2 disconfirmed the model's prediction that correct rejections to XY pairs should be faster than hits. To account for both the accuracy and latency findings in Experiment 2, a new dual-process model of associative recognition that combines the SAM-RAM retrieval mechanisms and the random walk mechanism in some latency models has been proposed. This new model is the first attempt to quantitatively describe both accuracy and latency data in associative recognition. With necessary further testing and development, it appears to be a promising one as it is providing decent predictions to the accuracy and latency data in Experiment 2. As mentioned above, it is potentially able to predict the similarity effect on RT data as shown in Experiment 2 and it needs to be fit to the real time, instead of units of time.

REFERENCES

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

Brown, S. & Heathcote, A. (2005). A ballistic model of choice response time. *Psychological Review, 112*, 117-128.

Busemeyer, J. R., & Townsend, J. T. (1992). Fundamental derivations from decision field theory. *Mathematical Social Sciences, 23,* 255–282.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100,* 432–459.

Criss, A.H. & Shiffrin, R.M. (in press). Pairs do not suffer interference from other types of pairs or single items in associative recognition. *Memory & Cognition.*

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

Diller, D.E., Nobel, P.A., & Shiffrin, R.M. (2001). An ARC-REM model for accuracy and response time in recognition and recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 27,* 414-435.

Dosher, B. A. (1984). Discriminating Preexperimental (Semantic) from Learned (Episodic) Associations: A Speed-Accuracy Study. *Cognitive Psychology, 16,* 519-555.

Egan, J. P. (1958). *Recognition memory and the operating characteristic.* Indiana University Hearing and Communications Laboratory, AFCRC-TN-58-51.

Francis, W. N. & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton Mifflin.

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

Greene, R. L. (2004). Recognition memory for pseudowords. *Journal of Memory and Language, 50,* 259-267.

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

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

Hinztman, D. L. & Curran, T. (1995). When encoding fails: Instructions, feedback, and registration without learning. *Memory & Cognition, 23,* 213-226.

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

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

Hockley, W. E. (1992). Item versus associative information: further comparisons of forgetting rates. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1321-1330.

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

Hockley, W. E. & Murdock, B. B. (1992). Speed-accuracy trade-off and item recognition: A reply to Gronlund and Ratcliff. *Journal of Mathematical Psychology, 36,* 461-467.

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

Kelley, R. & Wixted, J. T. (2001). On the nature of associative information in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 701-722.

Macmillan, N, A. & Creelman, C. D. (2004). Detection Theory: A User's Guide. Lawrence Erlbaum Associates, Inc.: New Jersey.

Malmberg, K. J., Holden, J. E., & Shiffrin, R. M. (2004). Modeling the effects of repetitions, similarity, and normative word frequency on judgments of frequency and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 319-331.

Malmberg, K. J. & Shiffrin, R. M. (in press). The "one-shot" hypothesis and the list-strength effect for free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

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

McClelland, J. L. & Chappell, M (1998). Familiarity breeds differentiation: A subjective-likelihood approach to the effects of experience in recognition memory, *Psychological Review, 105,* 724-760.

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

Murdock, B. B. (1993). TODAM 2: A model for the storage and retrieval of item, associative, and serial-order information. *Psychological Review, 100,* 183-203.

Murnane, K., Phelps, M. P., & Malmberg, K. (1999). ICE: A theory of context dependent discrimination. *Journal of Experimental Psychology: General, 128,* 403-415.

Pike, R. (1973). Response latency models for signal detection. *Psychological Review, 80,* 53-68.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59-108.

Ratcliff, R. & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review, 111,* 333-367.

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

Sheffert, S. M. & Shiffrin, R. M. (2003). Auditory registration without learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 10–21.

Shiffrin, R. M., Huber, D. E., & Marinelli, K. (1995). Effects of category length and strength on familiarity in recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 267-287.

Shiffrin, R.M., & Steyvers, M. (1997). A model for recognition memory: REM – retrieving effectively from memory. *Psychonomic Bulletin & Review, 4,* 145-166.

Shiffrin, R.M., & Steyvers, M. (1998). The effectiveness of retrieval from memory. In M. Oaksford & N. Chater (Eds.), *Rational models of cognition* (pp. 73-95). Oxford, England: Oxford University Press.

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes.* Cambridge, England: Cambridge University Press.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108,* 550–592.

Van Zandt, T., Colonius, H., & Proctor, R. W. (2000). A comparison of two reaction-time models applied to perceptual matching. *Psychonomic Bulletin and Review*, 7, 208-256.

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

Yonelinas, A.P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition, 25,* 1397-1410.

APPENDIX A. TABLES

Table 1

Means and standard error of measurement (SEMs) of Hit Rates and False-Alarm Rates for Experiment 1 (study time1.5 sec)

		HR			FAR		
Stimulus type		1	2	6	1	2	6
Words	Mean	.637	.771	.926	.255	.268	.184
	SEM	(.031)	(.031)	(.025)	(.031)	(.034)	(.037)
Pseudowords	Mean	.448	.620	.850	.325	.360	.476
	SEM	(.025)	(.025)	(.020)	(.025)	(.027)	(.030)
Characters	Mean	.427	.552	.699	.385	.480	.506
	SEM	(.025)	(.025)	(.020)	(.025)	(.027)	(.030)

Table 2

Means and SEMs of Hit Rates and False-Alarm Rates for Experiment 1 (study time 4 sec)

		HR			FAR		
Stimulus type		1	2	6	1	2	6
Words	Mean	.688	.857	.943	.190	.146	.143
	SEM	(.032)	(.028)	(.024)	(.032)	(.033)	(.037)
Pseudowords	Mean	.612	.740	.917	.375	.440	.388
	SEM	(.026)	(.022)	(.019)	(.026)	(.027)	(.030)
Characters	Mean	.555	.667	.832	.444	.442	.419
	SEM	(.026)	(.022)	(.019)	(.026)	(.027)	(.030)

Table 3

Means and SEMs of Reaction Time (ms) for Experiment 1 (study time 1.5 sec)

		Correct			Incorrect		
Stimulus type		1	2	6	1	2	6
Words							
Targets	Mean	2447	2191	1887	3543	3505	4087
	SEM	(157)	(117)	(119)	(294)	(305)	(711)
Foils	Mean	3074	3118	3018	3278	3063	2555
	SEM	(164)	(215)	(190)	(313)	(276)	(209)
Pseudo words							
Targets	Mean	2889	2639	2091	2892	2740	3317
	SEM	(160)	(124)	(100)	(173)	(129)	(359)
Foils	Mean	3064	3050	3398	3195	3138	2887
	SEM	(154)	(150)	(193)	(211)	(238)	(197)
Characters							
Targets	Mean	2892	2883	2626	3330	3298	3788
	SEM	(190)	(153)	(132)	(207)	(199)	(275)
Foils	Mean	3391	3057	3397	3271	3169	3169
	SEM	(173)	(198)	(234)	(246)	(272)	(217)

Table 4

Means and SEMs of Reaction Time (ms) for Experiment 1 (study time 4 sec)

		Correct			Incorrect		
Stimulus type		1	2	6	1	2	6
Words							
Targets	Mean	2327	2177	1818	3332	3884	4028
	SEM	(131)	(116)	(84)	(236)	(498)	(814)
Foils	Mean	2843	2868	2816	3584	3028	2707
	SEM	(157)	(147)	(135)	(427)	(248)	(316)
Pseudo words							
Targets	Mean	2716	2421	2030	3224	3523	3216
	SEM	(140)	(118)	(85)	(188)	(240)	(476)
Foils	Mean	3056	3312	3610	3247	2969	2765
	SEM	(160)	(202)	(177)	(232)	(171)	(216)
Characters							
Targets	Mean	3040	2868	2694	3348	3527	3614
	SEM	(165)	(169)	(132)	(206)	(258)	(426)
Foils	Mean	3325	3590	3849	3449	3481	3308
	SEM	(213)	(221)	(214)	(237)	(240)	(259)

Table 5

Means and SEMs of Hit Rates and False-Alarm Rates for Experiment2

		Presentations			
		0	1	2	6
AB (HR)	Mean		0.774	0.896	0.968
	SEM		(.022)	(.014)	(.008)
AD (FAR)	Mean		0.262	0.274	0.339
	SEM		(.041)	(.046)	(.055)
A'D' (FAR)	Mean		0.107	0.083	0.137
	SEM		(.031)	(.022)	(.029)
A'B' (FAR)	Mean		0.101	0.125	0.143
	SEM		(.027)	(.034)	(.036)
XY (FAR)	Mean	0.118			
	SEM	(.027)			

Table 6

Measn and SEMs of Reaction Time (ms) for Experiment 2

		Correct				Incorrect			
		0	1	2	6	0	1	2	6
AB	Mean		1825	1598	1398		2341	2212	2691
	SEM		(94)	(71)	(62)		(158)	(188)	(549)
AD	Mean		3086	2696	3229		2522	2488	2414
	SEM		(241)	(179)	(209)		(245)	(311)	(321)
A'D'	Mean		1848	2054	2093		2654	3768	2292
	SEM		(100)	(128)	(117)		(419)	(1017)	(298)
A'B'	Mean		2229	2352	2117		3104	2389	2577
	SEM		(162)	(184)	(152)		(703)	(409)	(346)
xy	Mean	1773				2421			
	SEM	(91)				(251)			

Notes: RT > (mean + 3 * SD) are removed

APPENDIX B. FIGURES

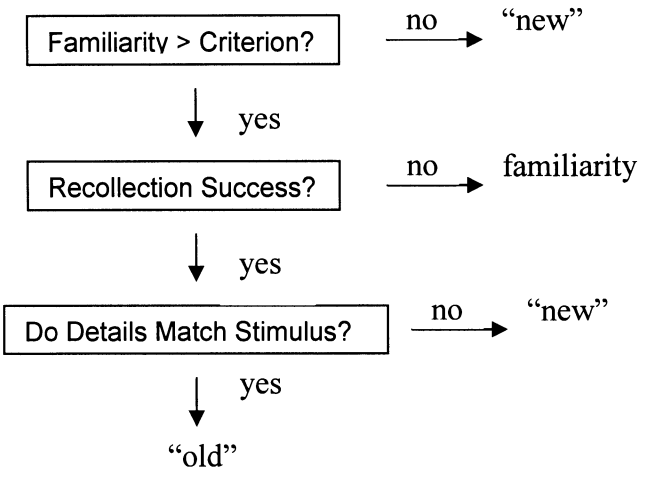

Figure 1. Flowchart for the REM Dual-Process Model.

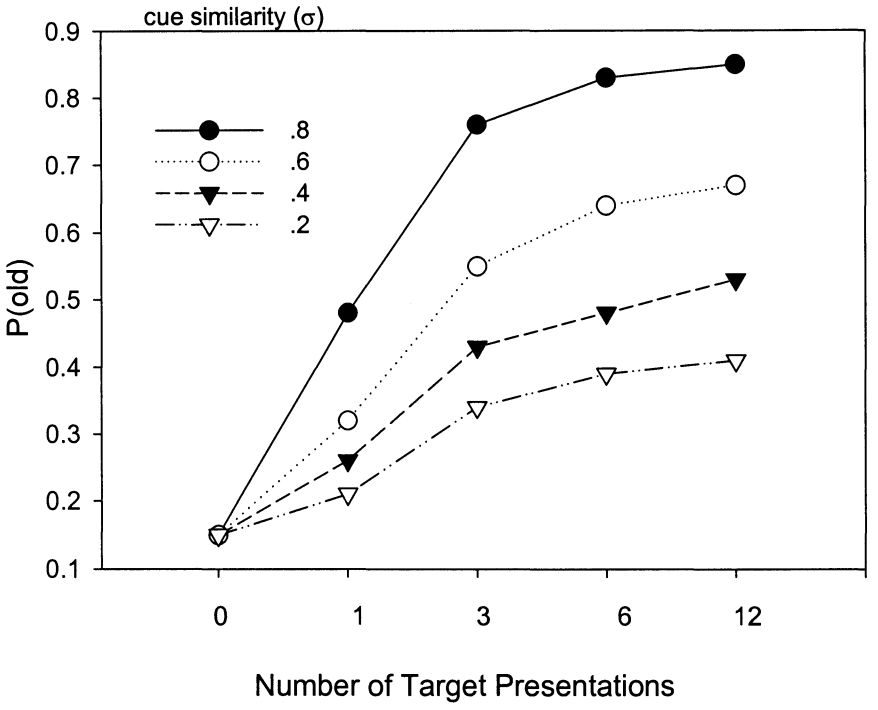

Figure 2. The Effect of Cue Similarity on FARs in the Single-Process Familiarity-Based REM Model. Parameter values: g = .4, w = 10, t = 1, u* = .5, c = .68, criterion = 1.0, v = .05.

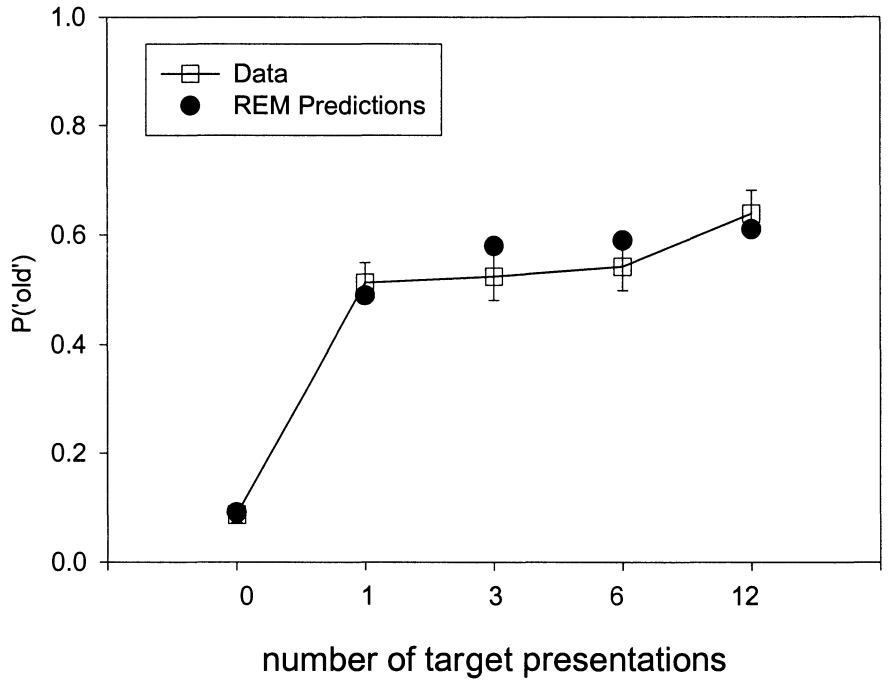

Figure 3. False-Alarm Rates as a Function of the Number of Presentations of a Target on the Study List.

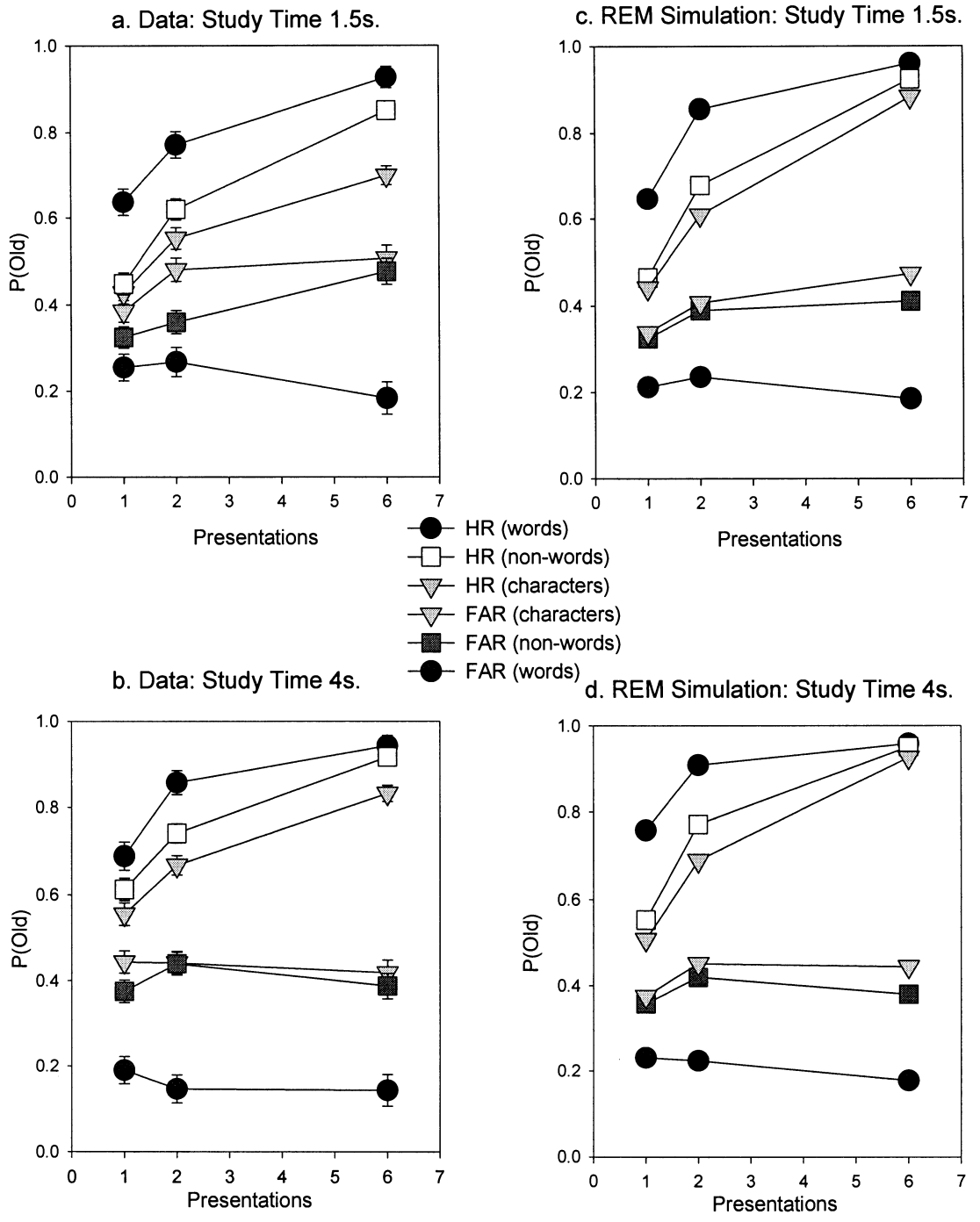

Figure 4. Mean HRs and FARs and REM Simulations for Experiment 1. Parameter values: $g = .4$, $c = .65$, $w = 20$, $t = 7.0$, $b = 1.0$; words: $u^* = .06$, $a = 1.0$, criterion1 = 1.0, criterion2 = 1.8; nonwords: $u^* = .03$, $a = .8$, criterion1 = .6, criterion2 = .8; characters: $u^* = .013$, $a = .8$, criterion1 = .6, criterion2 = .8.

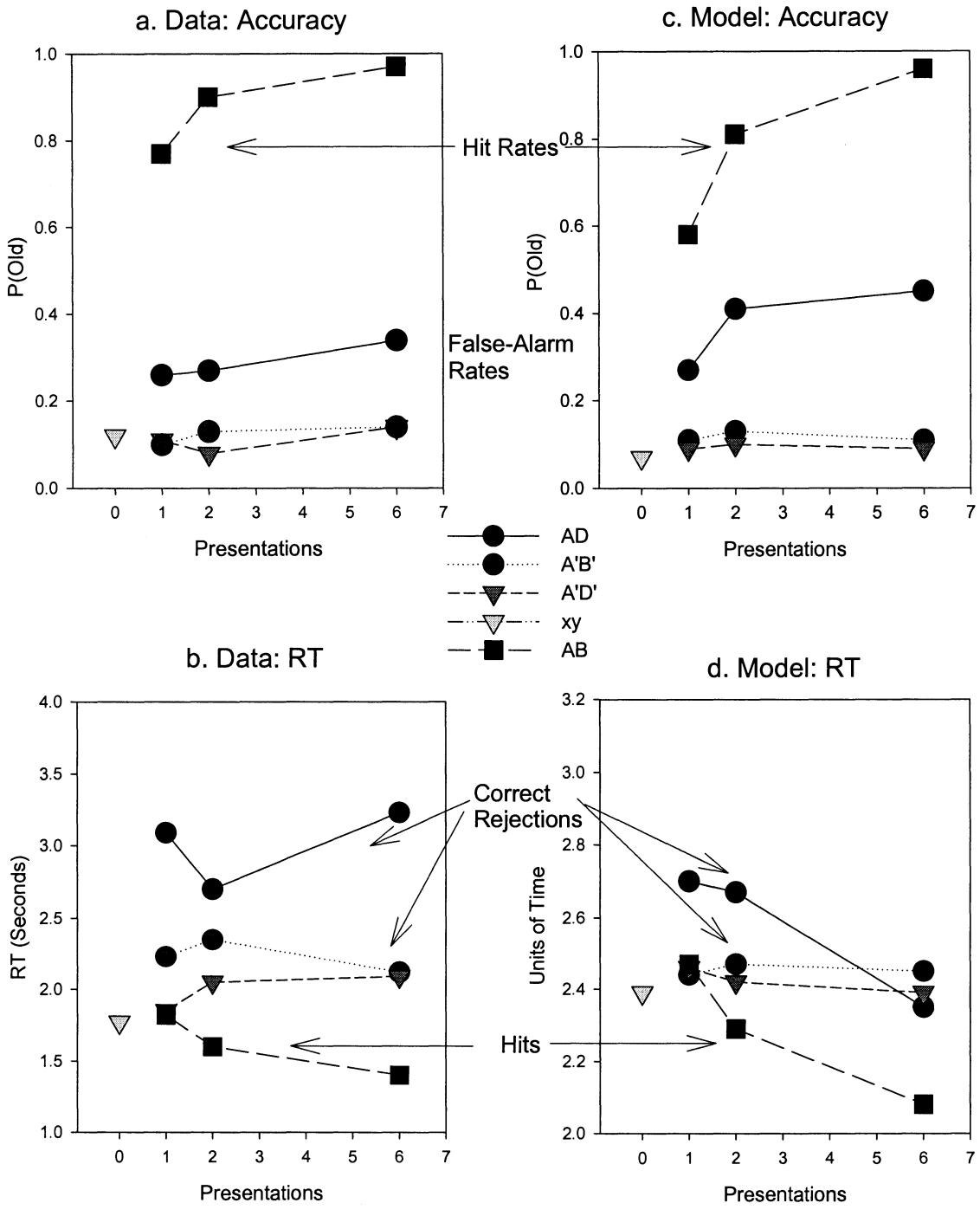

Figure 5. Mean Accuracy (HRs & FARs) and Latency (RTs) data and REM Simulations for Experiment 2. Parameter values: $g = .45$, $u^* = .04$, $c = .7$, $w = 20$, $t = 7$, $s = .26$, criterion1 = -0.4, criterion2 = 1.0, similarity = 0.30.

ACKNOWLEDGMENTS

I would like to take this opportunity to thank the individuals who have been helpful in the completion of this thesis and in my studies.

I would like to thank Dr. Kenneth Malmberg for his guidance in my research project and for his teaching of everything I know about being a serious scientist, a good modeler, and an active critical thinker, as well as all the ethical and technical details in conducting research. I also want to thank him for carefully revising my thesis draft to improve its quality.

I would also like to thank several teachers from whom I benefited a lot in my development. I want to thank Dr. Veronica Dark for her support and guidance in every stage of my research career and her thoughtful comments on my thesis proposal. I would also like to thank Dr. Vasant Honavar for his enlightening lecture, which had first directed my attention to cognition, and for his machine-learning class, which exposed me to the sophisticated technology that will help me in my future research. I would like to thank Dr's Alison Morris, Anne Cleary, Douglas Bonnet, Eric Cooper, and Gary Wells for teaching me the basic knowledge and research methods in cognitive psychology, which has not only broadened my vision, but also kept my strong interest in this fascinating field.

Several friends and my fellow graduate students have also contributed their support. I would like to thank Dan Peterson for collecting some of the data for me. I especially enjoyed the discussions with Moses Langley on various topics in memory research.

Lastly, I would like to thank Weiguo for being such a caring husband. I would not have achieved anything without his encouragement and support.