January 2007

# Negative Remembering

Aycan Kapucu
*University of Massachusetts Amherst*, akapucu@psych.umass.edu

# NEGATIVE REMEMBERING

A Thesis Presented

by

AYCAN KAPUCU

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE

September 2007

Psychology

NEGATIVE REMEMBERING

A Thesis Presented

by

AYCAN KAPUCU

Approved as to style and content by:

---

Caren M. Rotello, Chair

---

Neil A. Macmillan, Member

---

Thomas Kida, Member

---

Melinda A. Novak, Department Chair
Psychology Department

**ABSTRACT**

NEGATIVE REMEMBERING

SEPTEMBER 2007

AYCAN KAPUCU, B.A., BOGAZICI UNIVERSITY ISTANBUL

M.S., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Caren M. Rotello

Three experiments investigated the use of recall-to-accept and recall-to-reject processes in recognition and remember-know decisions. In all three experiments, participants studied a mixed list of singular and plural words. During the recognition test, participants made old-new confidence ratings and remember-know judgments for studied items, lures that were similar to studied items, and new lures. Old-similar ROC curves were constructed from the confidence ratings and found to be linear, consistent with the use of a high-threshold recollective process. The ROC intercepts and remember response rates converged on the same estimates of the amount of recollection for both positive (recall-to-accept) and negative (recall-to-reject) decisions.

CONTENTS

Page

ABSTRACT…………………………………………………………………………...iii

LIST OF TABLES……………………………………………………………………vi

LIST OF FIGURES…………………………………………………………………...vii

CHAPTER

1. INTRODUCTION…………………………………………………………………...1

ROC Analysis………………………………………………………………………..1
Remember-Know Paradigm…………………………………………………………7

Process-pure Interpretation……………………………………………………..8
One-dimensional Signal Detection Theory Interpretation………………...9

Problems with One-dimensional Model……………………………………………10

Response Bias……………………………………………………………………11
Two-point Slopes…………………………………………………………………...12
STREAK……………………………………………………………………………13

Combining ROC Analysis and R-K Paradigm…………………………………..14

2. EXPERIMENT 1……………………………………………………………………...16

Method…………………………………………………………………………………...16

Participants………………………………………………………………………….16
Stimuli……………………………………………………………………………….16
Design……………………………………………………………………………….16
Procedure…………………………………………………………………………...17
Analysis……………………………………………………………………………..18

Results…………………………………………………………………………………20
Discussion……………………………………………………………………………..22

3. EXPERIMENT 2……………………………………………………………………...25

Method……………………………………………………………………...25

Participants…………………………………………………………….25
Stimuli & Procedure…………………………………………………..25

Results……………………………………………………………………25
Discussion………………………………………………………………..27

4. EXPERIMENT 3……………………………………………………………...29

Method……………………………………………………………………...30

Participants…………………………………………………………….30
Stimuli………………………………………………………………….30
Procedure……………………………………………………………...30

Results……………………………………………………………………31
Discussion………………………………………………………………..33

5. GENERAL DISCUSSION…………………………………………………35

REFERENCES……………………………………………………………………..64

LIST OF TABLES

Table Page

1. Remember responses and intercepts for Experiment 1..................................................43

2. Remember responses and intercepts for Experiment 2..................................................44

3. Remember responses for Experiment 1 and Experiment 2............................................45

4. Remember responses and intercepts for Experiment 3 (With-New blocks)..................46

5. Remember responses and intercepts for Experiment 3 (No-New blocks).....................47

LIST OF FIGURES

Figure Page

1. Hypothetical ROC curves and z-ROC curves predicted by a continuous familiarity model that assumes equal-variance (a) and unequal variance (b)......................................48

2. Hypothetical ROC and z-ROC curves predicted by a high-threshold model................50

3. Hypothetical ROC and z-ROC curves predicted by a dual-process model..................51

4. Hypothetical ROC curve showing the use of recall-to-accept and recall-to-reject processes....................................................................................................................52

5. Signal-detection model for the remember-know paradigm...........................................53

6. The sum-difference theory of remembering and knowing............................................54

7. ROC curves for the old-new (a) and the old-similar (b) data in Experiment 1.............55

8. z-ROC curves for the old-new (a) and the old-similar (b) data in Experiment 1..........56

9. ROC curves for the old-new (a) and the old-similar (b) data in Experiment 2.............57

10. z-ROC curves for the old-new (a) and the old-similar (b) data Experiment 2............58

11. The old-similar ROC curves from Heathcote et al. (2006).........................................59

12. ROC (a) and z-ROC (b) curves for the old-new data in Experiment 3........................60

13. ROC curves for the old-similar data for With-New (a) and No-New (b) blocks in Experiment 3..............................................................................................................61

14. z-ROC curves for the old-similar data for With-New (a) and No-New (b) blocks in Experiment 3..............................................................................................................62

15. ROC curves for the old-similar data for B1-first (a) and B2-first (b) blocks.............63

# CHAPTER 1

# INTRODUCTION

The processes that underlie recognition memory performance have been extensively investigated. Despite the large body of work, it is still a question whether one or two processes are involved. For instance, we all probably have experienced meeting a person who feels familiar but about whom we recollect no details: not her name, where we met her before, or any other details. At other times, however, we immediately recognize the person and are able to identify who the person is. These two different components of the recognition experience are referred as *familiarity* and *recollection* in recognition memory.

It has been a matter of debate whether recognition judgments are based on either familiarity or recollection, or both. Single process models assume that recognition is based solely on the assessment of familiarity, they can be contrasted with dual process models in which recollection is assumed to accompany familiarity in making recognition decisions. Several experimental techniques have been used to obtain evidence for the use of recollection and familiarity components of recognition. For the purposes of the present study, I will discuss two of those techniques: *ROC analysis* and the *remember-know paradigm*.

## ROC Analysis

An important line of evidence for the use of recollective processes in recognition memory judgments comes from receiver operating characteristics (ROC) curves. One way of generating an ROC curve is by asking participants to rate their confidence that each item was studied (Macmillan & Creelman, 2005). Plotting the hit rate against the false-alarm rate at each confidence level yields an empirical ROC curve. ROC curves describe the relationship between hit and false-alarm rates at different levels of response bias or confidence. In a standard recognition memory task, for example, after studying a list of words, participants judge whether or not each test item had been studied and make their confidence response on a 6-point rating scale, ranging from “sure new” (1) to “sure old” (6). Theoretically, the lowest point (first pair of hit- and false-alarms) on the curve reflects the strictest response criterion that includes only the most confident old responses (i.e., “sure old”). The second lowest point reflects a slightly relaxed response criterion, which sums the proportion of the most confident old responses and the proportion of the second most confident old responses. The procedure is repeated until 5 (false alarm, hit) pairs are plotted. The last response category is not plotted because its hit and false-alarm rates are necessarily equal to 1.0. Therefore, a confidence scale with N ratings produces an ROC curve with N-1 points. Theoretically, all (false-alarm, hit) pairs on the curve represent the same sensitivity, differing only in terms of response bias.

The shape of the ROC curve provides information about the processes underlying recognition judgments. Different recognition memory models make different assumptions about these underlying processes and thus predict different ROCs. *Signal-detection* models assume that recognition decisions are based on a continuous familiarity dimension on which old items are more familiar than new items (Macmillan & Creelman, 2005). When the underlying item distributions have equal variance, the model predicts curvilinear and symmetric ROCs along the minor diagonal in probability space, with a slope 1.0 in z-space. However, when distributions have unequal variance and the old item distribution is more variable as is typically observed empirically, the ROC curve becomes asymmetric in probability space, with a slope less than 1 in z-space. When ROCs predicted by the signal-detection model are transformed into z-space, the resulting z-ROC curve is linear and the y-intercept of the z-ROC indicates sensitivity (*d´*) when the distributions are normal with equal variance (see Figure 1).

Another possibility is that recognition might be based on an all-or-none single high-threshold recollection process. Old items are recollected with high confidence and new items are never recalled regardless of how familiar they seem. “Old” responses to lures cannot exceed the threshold and therefore occur only due to guessing. The proportion of studied items that are called “old” by virtue of exceeding the high threshold creates a nonzero y-intercept for the ROC, other studied items may be guessed “old” and guess responses are randomly assigned to various levels of confidence. As a result, both hit and false alarm rates increase at the same pace for each confidence level, creating a constant slope. Thus, these high-threshold models predict a linear ROC with a y-intercept greater than zero and a curvilinear z-ROC (see Figure 2).

A mixed model in which both high-threshold recollection and familiarity contribute to recognition performance would predict ROCs with a y-intercept that is greater than zero, a curvilinear shape and a decreasing slope. For example, in the *dual-process* model proposed by Yonelinas (1994), a continuous, an equal-variance familiarity process (as in the signal-detection model) is accompanied by a high-threshold recollection process. The two processes operate independently in their contribution to recognition. ROC curves predicted by the dual-process model have a y-intercept greater than zero reflecting the proportion of recollected studied items and have an asymmetric curvilinear shape similar to that predicted by unequal-variance signal-detection model.

The model predicts a curvilinear z-ROC and the amount of curvature depends on the relative contribution of the recollective processes (see Figure 3).

Several recollective processes have been shown to contribute to recognition performance under different conditions (Rotello & Heit, 1999; Rotello, Macmillan, & Van Tassel, 2000). The acceptance of studied items might be facilitated by a *recall-to-accept* process in which evidence that matches the memory trace to the studied test probe is retrieved from memory. Alternatively, a *recall-to-reject* process facilitates the rejection of test lures that are similar to the studied items. In that process, mismatching information is retrieved from memory that helps to reject those similar lures. For an illustration of these processes, suppose that *computer* has been presented on the study list but *computers* has not. When participants are tested on *computer,* they are able to recall this item and correctly categorize it as 'old' by retrieving the matching memory trace. Alternatively, when they are tested on *computers,* they are able to correctly reject this item on the basis of a recall process that retrieves the mismatching memory trace of the studied item *computer.* Because participants know that *computer* and *computers* cannot both be on the study list, information about a recalled studied item is used to specifically to reject the other similar unstudied item which cannot be recalled. The rejection of completely new items (e.g., *table*) that are not similar to the studied items is not affected by this process because completely new items can be rejected on the basis of their low familiarity and cannot be rejected by recalling other words from the study list.

Over the time course of the recognition judgments, the early stage of processing is assumed to be determined by a familiarity-based process, whereas a recall-to-reject process begins to operate in a later stage when recall begins to contribute to the judgment

(Rotello & Heit, 1999). Consequently, the rejection of similar lures should increase once the information from a recall-to-reject process becomes available. Because the rejection of completely new lures is not affected by recall process, the false alarm rate to similar lures decreases more quickly than the false alarm rate to completely new lures due to the recall-to-reject process later in the time course.

The rejection of test lures that are similar to the studied items can be managed through a recall-to-reject process when the familiarity-based process is not helpful. Converging evidence for the use of recall-to-reject processing comes from different memory paradigms (Yonelinas, 1997; Rotello, Macmillan & Van Tassel, 2000). Yonelinas (1997) found evidence for recall-to-reject in an associative recognition paradigm. In that paradigm, participants study a list of word pairs (A-B, C-D) and later they are tested on intact (i.e., studied pairs: A-B) and rearranged pairs (i.e., each word of the pair has been studied but not paired together in the study list: A-D). The rearranged pairs are then similar lures that feel familiar despite being new items. The recall-to-reject process would operate because participants would correctly say 'new' to a rearranged pair since they recollected that the words were not paired in the study list (i.e., mismatching information).

Rotello et al. (2000) demonstrated that recall-to-reject can operate in item recognition as well. Subjects studied a mixed list of singular and plural words but never saw the same word in both forms. Later, they were tested on three types of items: studied words which were in the same form as they appeared on the study list, similar lures that were plurality-changed versions of studied words and completely new words that did not appear on the study list. Rotello et al. (2000) concluded that participants were able to correctly reject similar lures and call those items ‘new’.

The use of recall-to-reject processing has also been found in other memory tasks, such as the memory conjunction paradigm in which studied compound words (jailhouse; blackbird) are combined at test (jailbird) to form similar lures (e.g., Lampinen, Odegard, & Neuschatz, 2004; Jones, 2005). Lampinen et al. (2004) used ROC analyses to estimate recall-to-reject, finding it to be similar with the estimates obtained from associative recognition and item recognition tasks.

A recall-to-reject process shows itself on the ROC curve when similar lures are rejected with high confidence and no old items are given this high confidence “new” response. In this case, the ROC curve that plots the hit rate to old items against the false alarm rate to similar lures (i.e. *old-similar* ROC) intersects the upper x-axis where the false alarm rate is less than 1.0. The recall-to-reject probability can be estimated from the amount of reduction in this upper x-intercept. When the ROC is curvilinear with an upper x-intercept of 1, there is no evidence of recall-to-reject processing. As described earlier, ROCs are more linear to the extent that high-threshold recollection dominates the recognition decision. In that case, specific use of a recall-to-reject process is shown to produce a linear ROC with an upper x-intercept that is less than 1. Similarly, a recall-to-accept process that facilitates the recollection of studied items shows itself in ROC space with relatively linear functions that have a y-intercept that is greater than zero. Both recall-to-reject and recall-to-accept processes may operate simultaneously in recognition. Rotello et al. (2000) were able to show that the y-intercept of the ROC is informative about the presence of recall-to-accept processing and the upper x-intercept identifies the presence or absence of *recall-to-reject* processing in the recognition of similar lures (see Figure 4).

Other measures can also be used to estimate the independent contributions of familiarity and recollection-based processes to recognition judgments, such as multinomial modeling in the form of *conjoint recognition theory* (Branierd, Reyna, & Mojardin, 1999). Conjoint recognition theory provides an alternative to ROC analysis and employs a task which extends Jacoby's (1991) process dissociation procedure.[1]

Rotello (2001) combined conjoint recognition methodology with ROC analyses to estimate recall-to-reject and recall-to-accept processing. The goal was to compare the numerical estimates of recall-to-reject and recall-to-accept in two very different methodologies. Separate analyses of ROC data and fits of the conjoint recognition model were applied to the same data. The results provided converging evidence for the estimates of recall-to-accept and recall-to-reject because both methods provided similar estimates of those processes.

## Remember-Know Paradigm

We have used ROC analyses as one line of evidence for the use of recollective processes in recognition memory. Another line of evidence is the *remember-know* (R-K) paradigm. The paradigm was first suggested by Tulving (1985), who proposed that remembering and knowing are two different states of consciousness (i.e., autonoetic and noetic) which reflect two separate memory systems (i.e., episodic and semantic). The R-K paradigm has been widely explored and it has been shown that participants are capable of distinguishing between these two states of awareness. In a typical R-K study,

[1] The details of the process dissociation procedure are described in the original study by Jacoby (1991).

participants study some type of stimuli (i.e., words, pictures, etc…) and during the recognition test they are asked whether they studied each item. For items that are called "old", they are further asked to make a "remember" or "know" judgment. Participants are asked to give a "remember" response when they can recall specific details about the experience of studying that item, and to give a "know" response if they are aware having studied the item before but recall no details of the conscious experience of studying it.

Previous empirical research using the R-K paradigm can be grouped according to the assumptions made about the mechanisms that underlie R-K decisions. Here, I will briefly review two theoretical accounts with their major findings.

<u>Process-pure interpretation</u>

According to process-pure account, remember and know responses are direct measures of different underlying processes or memory systems. The variety of systematic dissociations between remember and know response rates as a result of using different experimental variables are taken as evidence for the judgments being process-pure.[2]

The idea that two separate memory systems underlie R-K judgments was first proposed by Tulving (1985), who stated that remembering reflected episodic memory whereas knowing reflected semantic memory. Later, following a similar kind of logic, Rajaram (1993) suggested that remember responses are affected by conceptual processing and know responses are affected by implicit processing as reflected in more perceptual tasks.

In the *dual-process model* (Yonelinas, 1994), which is a quantitative version of the process-pure account, remember and know are interpreted as process pure responses

---

[2] Major theories and empirical dissociations of R-K paradigm are described in a recent review by Gardiner and Richardson-Klavehn (2000).

that reflect recollection and familiarity, respectively (Yonelinas, 2001, 2002; Yonelinas et al., 1996). Furthermore, remember-know responses can be used to estimate the contribution of recollection and familiarity to recognition decisions. The assumption is that remember responses and ROC-based analyses should generate essentially identical estimates of these two underlying processes (Yonelinas, 2001; Yonelinas, Dobbins, Szymanski, Dhaliwal, & King, 1996). The dual-process model assumes a high-threshold recollection process and because remember judgments measure recollection, they should be diagnostic. Remember-false alarms should never occur because new items cannot exceed the high threshold. Therefore, remember false alarms are treated as occurring due to non-memorial noise in the data. For that reason, the amount of recollection can be estimated with the remember-hit rate minus the remember-false alarm rate. In other words, remember responses are regarded as high-confidence recognition decisions, corrected for guessing. Yonelinas (2001) showed that remember responses are made almost exclusively after highest confidence old judgments, and know responses are made across a variety of confidence levels.

One-dimensional signal detection theory interpretation

The standard signal detection theory (SDT) model assumes that recognition decisions are based on a unidimensional global memory strength value. Donaldson (1996) and Hirshman and Master (1997) applied such a signal detection model to data obtained from R-K judgments. According to this model, remember and know responses reflect different degrees of memory strength, or familiarity, rather than qualitatively different memory processes as suggested by the process-pure account. Recognition decisions are made through two separate criteria: the lower criterion (i.e., less stringent, liberal) is for making the recognition decision that an item is old or new, the higher criterion (i.e., more stringent, conservative) is used when making the R-K decision. Participants first respond old to the items whose familiarity exceeds the first criterion and the items that are called old are given a remember response if their familiarity exceed the second higher criterion (see Figure 5).

## Problems with one-dimensional model

There has been criticism of the one-dimensional SDT interpretation of R-K judgments. One criticism is that the signal detection model cannot explain why the decision criteria are influenced by different experimental variables (i.e., empirical dissociations) and that it does not suggest an explanation about how subjective states of awareness can be produced simply by a placement of different criteria on a single strength axis (Gardiner & Richardson-Klavehn, 2000). In response to these arguments, Dunn (2004) fit the one-dimensional model to the R-K data obtained from various experiments that had found dissociations and showed that the model was able to account for all of the data.

Moreover, Gardiner, Ramponi and Richardson-Klavehn (2002) proposed evidence against the model's basic prediction that memory sensitivity should remain constant when it is measured at the old-new criterion and R-K criterion. However, recent research (Macmillan, Rotello, & Verde, 2005; Dunn, 2004) has shown that the discrepancy between the estimates measured at two criteria is an expected consequence resulting from the natural properties of the measures (i.e., *A´* and *d´*) used to estimate sensitivity.

Response bias

The one-dimensional model predicts that manipulations that influence criterion placement (i.e., response bias) should influence both remember and know responses, in contrast to the dual-process account that predicts no change in remember response rates due to any changes in old-new decision criterion. Hirshman and Henzler (1998) investigated this prediction with an instructional manipulation: participants were told either 70% or 30% of the test items were old items (although there were actually 50% old and 50% new items in the test list). The results showed that participants used more liberal response criteria when they believed that test items were mostly studied items and more conservative criteria when they believed there were fewer studied items. Most importantly, criterion shifts affected both remember and know responses: liberal instructions increased both responses, whereas conservative instructions decreased them. Benjamin (2005) further explored criteria manipulations by using subjective distractor plausibility. Distractor plausibility was manipulated by presenting distractor items during the test either from the same or different semantic category with the studied items. Remember responses were increased by liberal shifts in criteria due to less plausible distractors.

In another investigation of changes in response bias for remember responses, Rotello, Macmillan, Reeder and Wong (2005) used an instructional manipulation regarding the definition of a remember response. Rotello et al. (2005) argued that the definition of what constituted a remember response was limited by the conservative instructions in Yonelinas's (2001) experiments therefore, the high-threshold nature of remember responses could merely be a consequence of response bias encouraged by the instructions. Participants in Yonelinas's (2001) study were told to give a remember response if they could explain the specific details of the recollection. However, in the standard R-K paradigm instructions any recollected aspect of the study experience could lead to a remember judgment (Rajaram, 1993). Using different instructions that manipulated participants' willingness to give a remember response, Rotello et al. (2005) were able to show that the remember response is subject to response bias and influenced by test instructions. When participants were given specific (narrow) instructions about what justifies a remember response as in Yonelinas's (2001) study, they produced fewer remember responses to words recognized with lower confidence. However, when the participants were given standard remember instructions (e.g., Rajaram, 1993), they reported remember responses over a range of confidence levels instead of only at the highest confidence.

Two-point slopes

Another criticism of the one-dimensional SDT interpretation was suggested by Rotello, Macmillan and Reeder (2004). They argued against the main assumption of the one-dimensional model that remember and old responses are merely high and low confidence decisions. If that were correct, then z-ROC slopes obtained from those two confidence levels should be equivalent to slopes obtained from simple confidence rating (old-new) data without R-K decisions. In a comprehensive meta-analysis, Rotello et al. (2004) computed z-ROC slopes from a two-point ROC curve from R-K data where the lower point represents "remember" hit and false alarm rates and the higher point represents "old" hit and false alarm rates. It turned out that the mean and variance of the two-point slopes obtained from old-new data and those obtained from R-K data were significantly different.

As a solution to this problem, Wixted and Stretch (2004) extended the one-dimensional model. They argued against the assumption of remember responses being high-confidence by proposing that variability in the R-K criterion (occurring with each test item) with respect to the old-new criterion would allow for a graded remember judgment. Thus, remember responses could occur at multiple confidence levels, rather than only in high-confidence levels.

STREAK

Taking into consideration the same problem about z-ROC slopes, Rotello et al. (2004) argued that R-K decisions might reflect some other form of evidence besides global memory strength and as a solution they developed a two-dimensional SDT model for the R-K paradigm. The sum-difference theory of remembering and knowing (STREAK) is fundamentally based on SDT framework but offers two strength dimensions on which recognition decisions are made (global familiarity and specific recollection), thus preserving the original idea of Tulving (1985) that remember and know responses reflect two separate forms of memory. According to the STREAK model (see Figure 6), recollection is a graded form of evidence like familiarity, and distributions of old and new items differ in their specific strength as well as their global strength. Old-new decisions are determined by the weighted sum of global and specific strengths, whereas R-K decisions are determined by the weighted difference of those two values. The two dimensions of the model are both continuous and the R-K bound is orthogonal to the old-new bound, so neither of the responses is process-pure. In conclusion, Rotello et al. (2004) argued that remember responses are not merely high-confidence old decisions and they do not reflect a high-threshold process. However, it is important to note for the purposes of the present study that Rotello et al. (2004) also stated that remember responses appear to result from a high-threshold process under particular experimental constraints (e.g., Rotello et al., 2005; Yonelinas, 2001).

In conclusion, there has not been an agreement in the recognition memory literature so far regarding the proper solution to the problems with the one-dimensional SDT interpretation of the R-K paradigm. Remember responses are argued to be graded, but they are also found to appear to be high-threshold recognition decisions in certain procedures. In other words, the exact nature of what it means to give a remember response has still not been resolved.

## Combining ROC analysis and R-K paradigm

We have examined evidence for two types of recollective processes (i.e., recall-to-accept and recall-to-reject) in recognition decisions. Converging evidence on the estimates of recall-to-reject and recall-to-accept has been obtained from both conjoint recognition methodology and ROC analyses. These estimates were obtained from the procedures that are predicted to produce high-threshold recollection. Previously, Rotello et al. (2000) showed that the y-intercept reflects recall-to-accept processing and the upper-x intercept reflects recall-to-reject processing. In addition, Yonelinas (2001) argued that remembering is diagnostic of recollection and thus the proportion of remember responses should be equal to the y-intercept which shows the proportion of items recollected. However, to our knowledge, the relationship between remember responses and recall-to-reject processing has not been investigated. The main purpose of the present study is to consider another way to estimate these two recollective processes.

The procedure that is used in the following experiments is exactly the same as that of Rotello et al. (2000), which previously yielded a high-threshold recollection process. In the present study, we extend that procedure by employing an R-K task in addition to simple item recognition in order to see whether remember responses in this narrowly constrained procedure will converge on the same values that have been obtained from both conjoint recognition and ROC analyses. In particular, we are interested in the question of how remember responses will correspond to the estimates of recall-to-accept and recall-to-reject processing.

# CHAPTER 2

# EXPERIMENT 1

Experiment 1 examined the hypothesis that remember responses, in both positive and negative senses, would converge on the same estimates obtained from ROC analyses, in particular the estimates of recall-to-accept and recall-to-reject processes. The procedure used in the present study was the same that of Rotello et al. (2000), except that it employed an R-K task in addition to an old-new confidence rating task. For that reason, the results of Rotello et al. (2000) were expected to be replicated. Accordingly, a curvilinear ROC curve was predicted for the old-new comparison and a more linear ROC curve was predicted for the old-similar comparison.

## Method

### Participants

Twenty-seven undergraduate students at the University of Massachusetts Amherst participated in exchange for a small payment or extra credit in their psychology courses. All participants were native English speakers.

### Stimuli

Three hundred common nouns (mean frequency, 70.6 per million; Kucera & Francis, 1967) and their plurals (e.g., *frog-frogs*) were selected from the MRC Psycholinguistic Database (Coltheart, 1981). Nouns were 3-12 characters in length and were chosen only if their plural form could be created by adding an “s”.

### Design

One word from each of 144 singular-plural pairs was pseudo-randomly selected for each subject to constrain that an equal number of singular/plural words were chosen.

The selected words were evenly divided into three study lists of 48 critical words each. Eight additional words were selected from the word pool for primacy and recency buffers for each list. Therefore, each study list was composed of 52 words in total. Presentation order was randomized for each list and each participant.

For each study list, a test list was created. The test list was composed of three types of items: Twenty-four of the studied words were selected at random to serve as Old items (twelve singular, twelve plural), the unstudied alternate forms (i.e., singular or plural) of the remaining twenty-four studied words served as the Similar lures, and an additional twenty-four completely unstudied words (twelve singular, twelve plural) were selected from the 144 singular-plural pairs to serve as New items. Test order was randomized for each subject.

In addition, a practice test list was constructed. It was composed of a total of six items: two Old words, two Similar lures and two New items. Two of the four primacy words were chosen to serve as Old items, the unstudied alternate form of the remaining two served as Similar lures and two unstudied words were chosen to serve as New items.

<u>Procedure</u>

Participants studied three lists of 52 words each, at a presentation rate of 3 ms. They were told to pay attention to the plurality of each word, in preparation for an unspecified memory test. Immediately following each list, participants received a recognition test for the words that had been studied previously (targets), the plural/singular alternate of the studied words (similar lures) and completely new words (new lures). For each test item, participants gave old/new confidence ratings on a 6-point scale (1 = sure new, 2 = probably new, 3 = maybe new, 4 = maybe old, 5 = probably old, 6 = sure old). Immediately following the old-new judgments, participants were asked to make a binary remember-know judgment for the same test item. The standard R-K instructions which were originally used by Rajaram (1993) were modified as follows: participants were asked to make a remember-know decision even if they had judged the test item to be “new”. Participants were told that word were presented in only their singular or plural form, and never both. They were also told that they could give a “remember” judgment if they remembered specific contradictory information, that is detail(s) that were inconsistent with the test item. They were also given the example that if they could recall studying *“computers”*, that meant that they could not have studied *“computer”* and so they could give *“computer”* a “sure new” judgment. Furthermore, they were told that if they had for instance remembered themselves imagining a computer lab full of computers while studying the word *computers,* they could still give a “remember” judgment to the test item *computer,* since that would imply a single computer, even though they had given a “sure new” judgment to *computer.*

<u>Analysis</u>

The ROC data in this experiment and the one that follows were analyzed in the following way. The probability of responding “old” to a studied item was compared to both to responding “old” to a completely new item and also to responding “old” to a similar item. In other words, two separate ROC curves were created in which the hit rate to old items was plotted against the false alarm rate to new items (i.e., *old-new* ROC) and the false alarm rate to similar items (i.e., *old-similar* ROC). In each case, a *group* ROC was generated by summing across the responses of the participants, and the Gaussian signal-detection model was applied to the group data. The fit of model to the data was done by maximum likelihood estimation, using SYSTAT's ROC module.

In order to test the high-threshold predictions for the *old-similar* ROC data (e.g., Rotello et al., 2000), a standard linear regression was conducted on the group *old-similar* ROC to determine whether there was a significant linear function. A quadratic term was then added to the linear function to determine whether it led to a significant improvement in the fit. To further assess the data, the group ROC was then replotted in *z*-space, and the same linear regression analyses were conducted to determine if the *z*-ROC was linear or U-shape. The threshold model predicts U-shaped z-ROCs, and the continuous model predicts linear *z*-ROCs. All fitting done with linear regression analyses was using least squared estimation. This method gives results that are essentially the same as those derived with maximum likelihood estimation (Ratcliff, McKoon & Tindall, 1994; Glanzer, Kim & Hilford, 1999).

In order to obtain estimates of recall-to-reject and recall-to-accept processes, the magnitude of the y-intercept and the upper x-intercept were estimated from the best-fitting regression line fit for the old-similar ROC data.

The Remember-Know data were analyzed the following way. First, remember responses that followed "old" recognition decisions were analyzed for both highest confidence (i.e., rating 6- "sure old") and also including lower confidence levels (i.e., ratings 4, 5 and 6). The probability of responding "remember" to an old item (i.e., remember-hit rate) and the probability of responding "remember" to a similar item (i.e., remember false-alarm rate) were computed. Then, a high-threshold correction was applied by subtracting remember-false alarm rate from remember-hit rate following high-threshold model assumptions (e.g., see Yonelinas et al., 1996).

The same procedure was repeated for remember responses following “new” recognition decisions, again for both highest confidence (i.e., rating 1- “sure new”) and lower confidence levels (i.e., ratings 1, 2 and 3). The high-threshold correction was applied this time by subtracting the remember rate to an old item from remember rate to a similar item.

The main issue is whether remember rates following “old” decisions converge on the recall-to-accept estimate (i.e., y-intercept) and remember rates following “new” decisions converge on the recall-to-reject estimate (i.e., upper x-intercept). For this reason, individual linear regression analyses were conducted for each subject to obtain individual estimates of intercepts. Moreover, corrected remember response rates were computed for each individual at both highest confidence and lower confidence levels. Then, paired-samples t-test was conducted on the corrected remember rates and intercept estimates to test whether they were statistically different from each other.

Results

The overall memory sensitivity for the present experiment and the one that follows is measured by $d_a$. The participants are able to discriminate old items from new items ($d_a = 1.105$) and they are also able to discriminate old items from similar items ($d_a = .756$).

The ROC data are shown in Figure 7. Best-fitting ROC curves are superimposed on both old-new and old-similar ROC data. The old-new ROC data are well fit by a continuous signal-detection model. However, the fit is less satisfactory for the old-similar

ROC data and the old-similar data are better fit by linear regression [quadratic component is not significant: t (3) = -1.282, p = .328]. The regression line superimposed on the data is shown in figure 7b. It has a y-intercept of .27 and an upper x-intercept of .78, indicating that 27% of the old items are recollected with high confidence by the use of a recall-to-accept process and 22% of the similar lures are rejected with high confidence by the use of a recall-to-reject process.

The old-new z-ROC and the old-similar z-ROC data are shown in Figure 8. The old-new z-ROC is linear in z-space with a small quadratic component ($zHit = .934 + .813\ zFA + .082\ zFA^2$). The old-similar z-ROC is best described by a quadratic function ($zHit = .674 + 1.220\ zFA + .324\ zFA^2$). The coefficients of the quadratic components of the old-new z-ROC and the old-similar z-ROC indicate that the curvature is more apparent in the old-similar z-ROC data.

The remember response rates are shown in Table 1. The probability of making a remember response after an "old" judgment is computed for both highest confidence "old" response and all "old" confidence levels and similarly, the probability of making a remember response after a "new" judgment is computed for both highest confidence "new" response and all "new" confidence levels. The crucial comparison between recollection estimates obtained from remember responses and from ROC data (i.e., y-intercept and upper x-intercept) reveals that the two methods do not converge for "old" recognition decisions [y-intercept = .27 vs. $P(R \mid \text{rating } 6) = .24$, $t(26) = 3.09$, p = .005 and y-intercept = .27 vs. $P(R \mid \text{ratings } 4,5,6) = .25$, $t(26) = 2.33$, p = .028] but they converge on the same estimates for "new" recognition decisions [upper x-intercept = .22 vs. $P(R \mid \text{rating } 1) = .20$, $t(26) = .98$, p = .334 and upper x-intercept = .22 vs.

*P* (R | ratings 1,2,3) = .22, *t*(26) = .27, p = .793].

## Discussion

The simple item recognition results of Rotello et al. (2000) were replicated. The old-new ROC data were curvilinear in the probability space and linear in the z-space. A continuous familiarity-based signal-detection model was able to account for the old-new ROC data. In such a signal-detection model, recognition decisions based on a continuous familiarity process produce curvilinear ROC curves (Macmillan & Creelman, 2005). However, the old-similar ROC data represented the opposite pattern, revealing a linear ROC in probability space and curvature in z-space. The continuous familiarity model did not appear to be capable of accounting for the old-similar data and instead, these results were more consistent with a high-threshold model of recognition memory. In the high-threshold model, recognition decisions are based on an all-or-none recollection process which produces linearity in the ROC curve. The old-similar ROC data were based on a recollection process because a familiarity-based process was not useful for discriminating studied words from similar test lures.

The results of Rotello et al. (2000) suggesting the use of specific recollective processes were also replicated. The old-similar ROC data reveal that both processes were operating. The y-intercept being greater than zero was indicative of a recall-to-accept process where in this case 27% of the studied items were recollected with high confidence and no similar lures were given this high-confidence “old” response. Similarly, the amount of reduction in the upper x-intercept in the old-similar ROC curve relative to the upper x-intercept of old-new ROC curve was indicative of a recall-to-reject process. Accordingly, 22% of the similar lures were rejected with high confidence and no studied words were given this high-confidence "new" response.

In the present study, the particular question we were interested in was the correspondence of remember responses to the estimates of recall-to-accept and recall-to-reject processes. Negative remember response rates converged on the same estimates of recall-to-reject, suggesting that remembering could be used in this negative sense as well as in the standard positive sense. However, positive remember responses reliably differed from the estimates of recall-to-accept (though numerically this difference was small). Previous research by Yonelinas (2001) showed that remember responses converged on the same estimates of recall-to-accept. On the other hand, it has recently been shown that the use of remember responses were highly influenced by the instructions provided to the participants (Rotello, Macmillan, Reeder, & Wong, 2005). Rotello et al. (2005) varied the remember instructions to investigate the effect of response bias. When participants were given more specific instructions about what justifies a remember response as they appeared in Yonelinas's (2001) study (such as they might need to explain their reasons for their remember responses or they were told to count the remembered details), they produced fewer remember responses to words recognized with lower confidence. Under these conservative instructions, recollection estimates obtained from ROC analyses and from remember judgments converged. However, when the participants were given standard remember instructions (e.g., Rajaram, 1993), they reported remember responses over a range of confidence levels and the estimates of recollection obtained from ROC analyses and remember responses differed.

In the present experiment, remember responses were spread over all levels of confidence. The participants produced remembering not only at the highest confidence level but at the lower confidence levels, as well. This result is not problematic for a continuous familiarity model where remembering is graded instead of being an all-or-none process. It may be problematic for a high-threshold model however, according to which remembering should occur only at the highest confidence level.

Although the procedure used in the current experiment was predicted to be constrained enough to lead participants to use high-threshold remembering, the possibility remains that maybe the instructions were not effective in leading the participants to use remember responses in the right way. Experiment 2 further examined this possibility by using more conservative remember instructions.

# CHAPTER 3

# EXPERIMENT 2

The procedure in Experiment 2 was exactly the same that of Experiment 1, with the exception of more conservative *remember* instructions being provided to the participants. The definition of remembering was intended to be even more constrained by instructing participants to make a remember judgment only if they could recall the specific plurality information. Fewer remember responses at the lower confidence levels were predicted and again remember responses were expected to converge on the estimates of recall-to-accept and recall-to-reject.

## Method

### Participants

Twenty-nine undergraduate students at the University of Massachusetts Amherst participated in Experiment 2 in exchange for a small payment or extra credit in their psychology courses. All participants were native English speakers.

### Stimuli & Procedure

The stimuli and the procedure were identical to those of Experiment 1, the only exception was that in Experiment 2, the definition of what constituted a "remember" judgment was narrowed. In other words, participants were asked to give a "remember" judgment only if they could recall the specific plurality information about the test item.

## Results

As in Experiment 1, the participants in Experiment 2 were able to discriminate old items from new items ($d_a$ = .880) and also old items from similar items ($d_a$ = .756).

The ROC data are shown in Figure 9. Best-fitting ROC curves are superimposed on both old-new and old-similar ROC data. The old-new ROC data were well fit by a continuous signal-detection model. However, the fit was less satisfactory for old-similar ROC data and the old-similar data were better fit by linear regression. The regression line superimposed on the data is shown in Figure 9b. It had a y-intercept of .25 and an upper x-intercept of .78, indicating that 25% of the old items were recollected with high confidence by the use of a recall-to-accept process and 22% of the similar lures were rejected with high confidence by the use of a recall-to-reject process.

The old-new z-ROC and the old-similar z-ROC data are shown in Figure 10. The old-new z-ROC was linear in z-space with a small quadratic component (zHit = .708 + .686 zFA + .054 $zFA^2$). The old-similar z-ROC was best described by a quadratic function (zHit = .720 + 1.162 zFA + .204 $zFA^2$). The coefficients of the quadratic components of the old-new z-ROC and the old-similar z-ROC indicated that the curvature was more apparent in the old-similar z-ROC data.

The remember response rates are shown in Table 2. The probability of making a remember response after an "old" judgment is computed for both highest confidence "old" response and all "old" confidence levels and similarly, the probability of making a remember response after a "new" judgment is computed for both highest confidence "new" response and all "new" confidence levels. Recollection estimates obtained from the remember responses and from the ROC data (i.e., y-intercept and upper x-intercept) converged for "old" recognition decisions [y-intercept = .25 vs. $P$ (R | rating 6) = .22, $t(28) = 1.47$, p = .153 and y-intercept = .25 vs. $P$ (R | ratings 4,5,6) = .25, $t(28) = .116$, p = .909] and the two methods also converged on the same estimates for "new" recognition decisions [upper x-intercept = .22 vs. $P$ (R | rating 1) = .21, $t(28) = .80$, p = .431 and upper x-intercept = .22 vs. $P$ (R | ratings 1,2,3) = .22, $t(28) = .36$, p = .722].

The remember rates for both Experiment 1 and Experiment 2 are presented together in Table 3. In Experiment 2, remember responses after "old" recognition decisions to both old items and similar lures decreased relative to Experiment 1. However, the participants still produced remember responses at all confidence levels instead of only at the highest confidence levels.

Discussion

As in Experiment 1, Experiment 2 replicated the item recognition results of Rotello et al. (2000). The old-new ROC data were well accounted for by a continuous familiarity model, whereas the old-similar data were better accounted for by a high-threshold model. Consequently, the old-similar ROC was linear with a nonzero y-intercept, due to a recall-to-accept process and a reduced upper x-intercept due to a recall-to-reject process.

Regarding the crucial comparison of the recollection estimates obtained from both the ROC analyses and the remember responses, the results of the present experiment confirmed our predictions that the two methods would converge on the same values. In particular, negative remember responses that were made after "new" recognition decisions converged on the estimates of recall-to-reject processing as it was the case in Experiment 1. Moreover, positive remember responses that were made after "old" recognition decisions also converged on the estimates of recall-to-accept processing in this experiment, contrary to the small but reliable difference observed between the two methods in Experiment 1.

There was an overall decline in the use of remembering for both old items and similar lures in Experiment 2 compared to Experiment 1, as a result of more conservative R-K instructions. However, the conservative instructions did not result in participants using remembering only at highest confidence levels. According to the high-threshold model assumptions, remembering operates on an all-or-none recollection process and therefore remember responses can truly be made only with high confidence. The observation of remember responses also at lower confidence levels seems to be in favor of the previous research suggesting that remembering is graded rather than having a threshold character (Rotello et al., 2005).

To sum up the item recognition results, Experiments 1 and 2 found a curvilinear ROC for the old-new data accounted for a continuous familiarity model and a linear ROC for the old-similar data accounted for a high-threshold model for both of the experiments. Therefore, it can be concluded that the results of Rotello et al. (2000) were successfully replicated.

CHAPTER 4

EXPERIMENT 3

Recently Heathcote, Raymond, and Dunn (2006) demonstrated a different pattern in their attempt to replicate Rotello et al.'s (2000) study. The researchers extended Rotello et al.'s (2000) procedure to employ an additional condition in which completely new lures are omitted from the test list. In other words, the participants are tested only on studied items and similar lures in this condition. This manipulation is based on the assumption that the presence of new lures encourages the use of familiarity in recognition decisions because new lures can be discriminated from studied items on the basis of familiarity. So, elimination of new lures from the test list could reduce the reliance on familiarity when discriminating similar lures from the studied items and lead to more linear old-similar ROCs. However, the results of Heathcote et al. (2006) suggested that the old-similar ROC data became even *less* linear when the new lures were removed from the test list. It is important to note that the old-similar ROC data was still less curved than the old-new ROC data, thus a recollection process must still be operating to a certain extent. However, the data are not accounted for by a high-threshold model, in contrast to the old-similar data of Rotello et al. (2000) and of the Experiments 1 and 2. Furthermore, the overall memory sensitivity was higher in the No-New condition compared to the condition in which new items were present (see Figure 11).

Considering the discrepancy in the old-similar ROC results of these separate studies, the present experiment aimed to replicate the results of Heathcote et al. (2006) in Experiment 3 by employing a modified version of the procedure used in their study. In order to explore the discrepancy between testing conditions either with or without new items, a within-subjects design was used in which recognition memory tests either including or not including new items will be presented in successive blocks. In particular, we wanted to see if we could replicate the reduced linearity and higher memory sensitivity in the condition without completely new items in Heathcote et al. (2006).

## Method

### Participants

Twenty undergraduate students at the University of Massachusetts Amherst participated in Experiment 3 for extra credit in their psychology courses. All participants were native English speakers.

### Stimuli

The stimuli were identical to those of Experiment 1 and Experiment 2.

### Procedure

The procedure remained the same to those of Experiments 1 and 2, except that in the present experiment there were six study-test blocks in which completely new items were either omitted or not omitted from the test. In both *With-New* and *No-New* blocks, the participants studied 24 singular and 24 plural items. For the memory test in With-New blocks, sixteen of the studied words were selected at random to serve as Old items (eight singular, eight plural), the unstudied alternate forms (i.e., singular or plural) of the sixteen studied words served as the Similar lures, and an additional sixteen completely unstudied words (eight singular, eight plural) were selected from the pool of singular-plural pairs to serve as New items. In No-New blocks, participants were only tested on sixteen Old items and thirty-two Similar items and the test did not include any completely New items. During the experiment, six study-test blocks were presented in a fashion that each With-New block was followed by a No-New block, or vice versa. Half of the participants started the experiment with a With-New block and the remaining half started with a No-New block.

Moreover, in the present experiment standard R-K instructions (as in Experiment 1) were used instead of the conservative instructions because there was not a substantial difference between the results of the two experiments due to different instructions. For each test item, the participants were asked to make a confidence rating and then make a binary remember-know judgment.

## Results

The *old-new* ROC data and *old-similar* ROC data from With-New and No-New blocks were analyzed separately. First of all, the overall memory performance measured by $d_a$ did not differ between the old-new and old-similar data and more importantly between With-New and No-New blocks. For With-New blocks, participants were able to discriminate Old items from New items ($d_a$ = 1.48) and also Old items from Similar lures ($d_a$ = 1.46). Similarly for No-New blocks, participants were able to discriminate Old items from Similar lures ($d_a$ = 1.26). The old-similar memory sensitivity did not differ between With-New blocks and No-New blocks [t(19) = 1.836, p = .08].

The old-new ROC and z-ROC data for With-New blocks are shown in Figure 12. The old-new ROC data were well fit by a continuous unequal-variance signal-detection model, as in Experiments 1 and 2. The old-new z-ROC data were well fit by a linear regression line (zHit = 1.206 + .634 zFA + .041 $zFA^2$) with a non-significant quadratic component (p = .126) confirming the continuous ROC data.

The old-similar ROC data for both With-New and No-New blocks are shown in Figure 13 with the best-fitting ROC curve superimposed on the data. Overall, the fit of the continuous signal-detection model was less satisfactory for both With-New and No-New blocks compared to the old-new ROC data. The With-New old-similar data were more linear than the continuous model would predict and better fit by a regression line (Hit = .368 + 1.78 FA – 1.273 $FA^2$) with a small quadratic component (p = .037). The No-New old-similar data were even more linear and better fit by a regression line (Hit = .409 + 1.381 FA - .829 $FA^{2)}$ with no significant quadratic component (p = .071).

The old-similar z-ROC data for both With-New and No-New blocks are shown in Figure 14. The old-similar z-ROC for With-New blocks was linear in z-space (zHit = 1.543 + 1.615 zFA + .374 $zFA^2$) with a small quadratic component (p = .037) and the old-similar z-ROC for No-New blocks was essentially linear (zHit = 1.241 + 1.226 zFA + .257 $zFA^2$) with a marginally significant quadratic component (p = .06).

The remember response rates for With-New and No-New blocks are shown in Table 4. Following the same procedure used in Experiments 1 and 2, the probability of making a remember response after an "old" judgment was computed for both highest confidence "old" response and all "old" confidence levels and similarly, the probability of making a remember response after a "new" judgment was computed for both highest confidence "new" response and all "new" confidence levels.

For With-New blocks, recollection estimates obtained from the remember responses and the old-similar ROC data (i.e., y-intercept and upper x-intercept) converged for "old" recognition decisions [y-intercept = .47 vs. *P* (R | rating 6) = .40, *t*(19) = 1.56, p = .135 and y-intercept = .47 vs. *P* (R | ratings 4,5,6) = .41, *t*(19) = 1.246, p

= .228]  and the two methods also converged on the same estimates for "new" recognition decisions [upper x-intercept = .46 vs. *P* (R | rating 1) = .41, *t*(19) = 1.672, p = .111 and upper x-intercept = .46 vs. *P* (R | ratings 1,2,3) = .42, *t*(19) = 1.277, p = .217].

For No-New blocks, recollection estimates obtained from the remember responses and the old-similar ROC data did not converge for the highest confident "old" recognition decisions  [y-intercept = .48 vs. *P* (R | rating 6) = .39, *t*(19) = 2.263, p = .036] but they converged for all "old" decisions combined [y-intercept = .48 vs. *P* (R | ratings 4,5,6) = .41, *t*(19) = 1.682, p = .109]  and the two methods also converged on the same estimates for "new" recognition decisions [upper x-intercept = .35 vs. *P* (R | rating 1) = .35, *t*(19) = -.009, p = .993 and upper x-intercept = .35 vs. *P* (R | ratings 1,2,3) = .38, *t*(19) = -.461, p = .650].

Recall that about half of the participants (N = 9) started the experiment with a With-New block and the remaining half (N = 11) started with a No-New block. In order to eliminate any possible confounding effects of the alternating With-New and No-New blocks, we conducted a separate between-subjects analyses on the first blocks only. The memory sensitivity did not differ between the No-New first block ($d_a$ = 1.34) and the With-New first block ($d_a$ = 1.20), [t (18) = -.370, p = .715].  The old-similar ROC data for the With-New first and No-New first blocks are shown in Figure 15. The data in both conditions could not be accounted for a one-dimensional signal-detection model and the linearity was apparent even in the first blocks.

Discussion

Overall, the results of Experiment 3 were in agreement with those of Experiments 1 and 2.  First, the old-new data produced curvilinear ROC curves that were well fit by the one-dimensional signal-detection model assuming continuous familiarity. Second, the old-similar ROC data could not be as successfully accounted for by the continuous familiarity model and produced more linear ROCs that could be best fit by linear regression.

The motivation for the present experiment was to replicate the results of Heathcote et al. (2006) using a within-subjects design, in order to be able to resolve the discrepancy between their results and the results of Rotello et al. (2000), as well as those of Experiments 1 and 2. In a between-subjects design, Heathcote et al. (2006) found that memory sensitivity was higher and ROCs were less linear and more spread out in the probability space in the condition without new test lures (i.e., No-New) compared to the condition including new lures (i.e., With-New). In the present experiment, we were unable to replicate these results (Heathcote et al., 2006). First, there were no significant sensitivity differences between With-New and No-New blocks, and even the sensitivity was slightly higher in the With-New condition. Although the first-block analyses of each condition suggested higher sensitivity when the first block was a No-New block, this superiority disappeared later on in the experiment. Second, in contrast to the reduced linearity in the No-New condition suggested by Heathcote et al. (2006), the results of the present experiment revealed that ROCs were even more linear in the No-New blocks compared to the With-New blocks.

# CHAPTER 5

## GENERAL DISCUSSION

Three experiments investigated two types of recollective processes (i.e., recall-to-accept and recall-to-reject) in recognition decisions using two different methods: ROC analyses and Remember-Know paradigm.

The one-dimensional signal detection model assumes that recognition decisions are based on a continuous familiarity process. In such a model, the ROC curve that plots the hit rate to Old items against the false alarm rate to New lures is curvilinear in probability space and linear in z-space. In contrast, a high-threshold model assumes that recognition decisions are based on a discrete recollection process and produces linear ROC curves with emerging intercepts.

The plurals paradigm has been previously shown to be one of the exceptional situations that produce high-threshold recollection in recognition decisions (Rotello et al., 2000). In such a paradigm, participants first study a mixed list of singular and plural common nouns and then perform a recognition memory test that includes three types of items: Old items (items studied on the previous study list), Similar lures (plurality-changed versions of the studied items) and New lures (completely new items never presented on the study list in either form). Rotello et al. (2000) showed that the old-similar ROC curve which plotted the hit rate to Old items against the false-alarm rate to Similar lures was linear in ROC-space, with a non-zero y-intercept reflecting recall-to-accept processing and a reduced upper x-intercept reflecting recall-to-reject processing. However, using the plurals paradigm, Heathcote et al. (2006) recently found that the old-similar ROC data were less linear than the high-threshold model would predict when the new lures were omitted from the test list.

All three experiments in the present study employed some modified version of the plurals paradigm used in Rotello et al. (2000), with the addition of R-K judgments following Old-New decisions. For each test item, the participants made an Old-New recognition judgment on a six-point confidence rating scale, followed by a binary R-K decision. Importantly, participants were asked to make an R-K judgment after each test item even though they had classified the item as “new”. Remember judgments that would follow a “new” recognition decision are defined as recollection of specific contradictory information about that test item. Recently, Jones (2005) has also asked participants to use negative remembering in a memory conjunction paradigm.

In all three experiments, the old-new ROC curves were curvilinear that and well-fit by the one-dimensional, continuous familiarity model, with z-ROCs confirming the model’s fit. However, the one-dimensional signal detection model failed to account for the old-similar ROC data. The data were more linear in probability space and had curvature in z-space which could be better accounted for by a high-threshold recollection process and the emergent intercepts provided estimates of recollection.

The critical question in the present study was whether the recollection estimates obtained from ROC intercepts and remember responses would converge on the same values. Previously, Yonelinas (2001) showed that old-new ROC intercepts and remember responses provided identical estimates of recollection in an associative recognition paradigm. In all three experiments, the two values generally agreed for “old” recognition decisions, in other words recall-to-accept estimates corresponded well to remember responses to the Old items. More importantly, the two methods always converged on the same estimates for "new" recognition decisions: remember responses to Similar items corresponded well to the estimates of recall-to-reject processing. This new finding leads to the conclusion that negative remembering as well as the positive remembering arises primarily from recollective processes in this particular paradigm. The plurals paradigm is a narrowly constrained case because the discrimination between the targets and lures is based only on the plurality information.

There are also a number of differences between the results of the three experiments. There was an overall decline in the rate of remembering in Experiment 2, due to the introduction of more conservative instructions. Instead of the standard R-K instructions (Rajaram, 1993), participants were asked to make a remember judgment only if they recollected the specific plurality information. The conservative instructions constrained the meaning of remembering and led to lower remember response rates. Indeed, Rotello et al. (2005) recently showed that remember responses were highly affected by instructions and they converge with ROC estimates only under conservative instructions that restricted use of remembering to the recollection of a single detail. The conservative instructions did not push remembering to the highest confidence levels only; instead, remember responses were spread over all confidence levels. According to the high-threshold model assumptions, remembering operates on an all-or-none recollection process and therefore remember responses can truly be made with high confidence. The observation of remember responses at lower confidence levels even under the conservative instructions seems to be in favor of the previous research suggesting that remembering is graded rather than having a threshold character (Rotello et al., 2005).

Overall performance was better in Experiment 3 than in Experiments 1 and 2. Memory sensitivity was higher, as well as the estimates of recall-to-accept and recall-to-reject, and there was also a significant increase in the rate of remembering. However, since the increase in the ROC intercepts was matched with the increase in the remember rates the two methods still converged on the same recollection estimates, confirming our primary prediction. There is no definite explanation for the increase in performance. Although the study lists were of the same length in all three experiments, the test lists in Experiment 3 were relatively shorter (i.e., 48 test items) than in Experiments 1 and 2 (i.e., 72 test items), therefore the shorter average study-test delay might be one of the possible explanations for higher performance.

The results of the current experiments are in disagreement with those of Heathcote et al. (2006). The higher memory sensitivity and reduced linearity of the old-similar ROC in the No-New condition were not replicated. There are a number of possible explanations for this discrepancy. Andrew Heathcote suggested (in a personal communication) that the discrepancy might be arising from the difference in the use of the confidence rating scale, in that the participants in the present study might have overused the highest confidence rating categories and failed to use the middle response categories. In other words, the linear ROCs might have resulted from the failure of participants to distribute their responses over the full range of the rating scale. Following this suggestion, I reanalyzed the data of Experiment 3 separately for two groups of participants who used and failed to use the middle response categories (i.e., 3 – "maybe new” and 4 – “maybe old”) for both With-New and No-New conditions. However, there were no apparent differences in the form of the ROC (i.e., linearity) between the two groups.

Another possible explanation is the within-subjects design of Experiment 3, which also was the only major difference from the design of Heathcote et al. (2006). Alternating presentations of With-New and No-New blocks might have affected the decision strategy of the participants, so I conducted a block-by-block analysis in order to see if the initial block would yield different results that were in the direction of Heathcote et al’s (2006). However, the ROCs observed in each of the No-New blocks were still more linear than the With-New blocks and the ROCs in both conditions were not well-fit by a Gaussian model.

Although the current experiments are in disagreement with Heathcote et al. (2006), they are along the similar lines with studies that previously suggested the use of a high-threshold recollection process in recognition memory with a variety of experimental paradigms. The plurals paradigm used in the current study has been previously demonstrated to generate linear ROCs in discriminating targets from similar lures (Rotello et al., 2000; Rotello, 2001). Whereas the discrimination between targets and new lures could be made on the basis of familiarity alone, the discrimination between targets and similar lures required recollective processing since both types of items were equally familiar.

Other linear ROCs in recognition memory come from source memory, associative recognition, and memory conjunction paradigms. Source memory judgments require recollection of qualitative information about the study episode and have been shown to produce linear ROCs (e.g., Rotello & Heit, 1999; Yonelinas et al., 1999). Similarly, in both associative recognition and memory conjunction paradigms, the critical discrimination between intact and rearranged item pairs entails specific recollection, rather than just familiarity, of the individual items constituting the pair. Therefore, these tasks result in linear ROCs (e.g., Yonelinas, 1997; Rotello et al., 2000; Lampinen et al., 2004).

It seems that the use of a high-threshold recollective processing in recognition memory depends heavily on the experimental task being narrowly constructed so that familiarity alone is not sufficient to discriminate targets from lures that differ from each other only on a single, specific detail (i.e., plurality, source, associative information). Relying on that single feature leads to discrete recollective processing which can be modeled by a high-threshold process instead of a Gaussian continuous familiarity process. These narrow tasks can be conceived as being special cases in which a threshold process appears to be used. In other words, the use of a discrete recollection process may be more like a strategy bias created by specific instructions and/or constrained tasks than a genuine basis for recognition memory decisions.

The debate of whether recognition decisions are based either on familiarity or recollection, or both, has not yet been resolved. The one-dimensional signal-detection model assumes a continuous familiarity process producing curvilinear ROCs, whereas high-threshold models assume a discrete recollection process. Linear ROCs, especially in combination with the R-K data presented in three experiments in this study support this king of recollection processing that is high-threshold in nature.

Models that incorporate both processes in recognition decisions differ in their assumptions on the nature of the recollection process and also on how the two processes interact with each other in their contribution in recognition. The dual-process model proposed by Yonelinas (1994) is a mixed model in which recollection is a high-threshold process and recognition decisions on a single trial are based on either familiarity or recollection, independently of each other. Dual-process model produces curvilinear ROCs due to the continuous familiarity process with a non-zero y-intercept due to high-threshold recollection. However, there are also mixed models in which recognition decisions are based on both processes, none of which is process-pure and threshold-like.

Recently, Wixted (2007; Wixted & Stretch, 2004) has proposed such a hybrid model framed in signal-detection theory that incorporates both familiarity and recollection as continuous processes, combined into a single dimension of memory strength. More specifically, recognition decisions are based on the information provided by the relative contributions of both processes. Wixted's model (2007) differs from the dual-process model proposed by Yonelinas (1994) in its assumptions of a continuous recollection process instead of a discrete high-threshold recollection. STREAK is also another two-dimensional model that assumes both recognition and R-K decisions are based on the weighted contributions of global familiarity and specific recollection, both of which reflect continuous memory strength (Rotello et al., 2004). Both STREAK and Wixted's hybrid model produce curvilinear ROCs due to the continuous underlying processes.

The data reported by many studies in recognition memory have made it quite clear that pure signal detection models based on a single process are not sufficient accounts of recognition performance. Both familiarity and recollection are graded, continuous processes and together contribute to recognition decisions. However, it also appears to be possible that high-threshold recollection might operate under certain conditions and experimental paradigms such as the plurals paradigm employed in the current study. It is still an open question that why exactly this particular paradigm leads to high-threshold processing and its underlying mechanisms must be investigated further by future research.

**Table 1. Remember responses (with means and standard errors) and intercepts for Experiment 1.**

| | "Old" | | | | "New" | | | |
|---|---|---|---|---|---|---|---|---|
| | High-confidence level (6) | | All levels (4,5,6) | | High-confidence level (1) | | All levels (1,2,3) | |
| | M | SE | M | SE | M | SE | M | SE |
| P (R \|old item) | .36 | .04 | .48 | .03 | .03 | .00 | .08 | .02 |
| P (R \|similar item) | .12 | .02 | .22 | .03 | .23 | .04 | .30 | .05 |
| P (R \|new item) | .03 | .01 | .09 | .02 | .11 | .02 | .18 | .03 |
| Threshold correction | .24[a] | .04 | .25[a] | .05 | .20[b] | .04 | .22[b] | .05 |

| | M | SE | M | SE |
|---|---|---|---|---|
| ROC estimates | .27[c] | .05 | .22[d] | .05 |

Notes: 1. Threshold corrections for (a) and (b) are performed as follows:

a. P (R | old item called "old" ) – P (R | similar item called "old")
b. P (R | similar item called "new") – P (R | old item called "new")

2. The ROC-based estimates of y-intercept (c) and upper x-intercept (d)

3. *P (R | old item)*: Probability of making a "remember" response to an "old" item.

**Table 2. Remember responses (with means and standard errors) and intercepts for Experiment 2.**

| | "Old" High-confidence level (6) | | "Old" All levels (4,5,6) | | "New" High-confidence level (1) | | "New" All levels (1,2,3) | |
|---|---|---|---|---|---|---|---|---|
| | M | SE | M | SE | M | SE | M | SE |
| P (R \|old item) | .30 | .04 | .41 | .04 | .05 | .01 | .11 | .02 |
| P (R \|similar item) | .09 | .01 | .17 | .02 | .25 | .04 | .32 | .04 |
| P (R \|new item) | .02 | .01 | .10 | .02 | .06 | .01 | .13 | .03 |
| Threshold correction | .22[a] | .04 | .25[a] | .05 | .21[b] | .04 | .22[b] | .05 |

| | "Old" M | "Old" SE | "New" M | "New" SE |
|---|---|---|---|---|
| ROC estimates | .25[c] | .04 | .22[d] | .04 |

Notes: 1. Threshold corrections for (a) and (b) are performed as follows:

c. P (R | old item called "old" ) – P (R | similar item called "old")
d. P (R | similar item called "new") – P (R | old item called "new")

2. The ROC-based estimates of y-intercept (c) and upper x-intercept (d)
3. *P (R | old item)*: Probability of making a "remember" response to an "old" item

**Table 3. Remember responses (with means) for Experiment 1 and Experiment 2.**

| | "Old" | | | | "New" | | | |
|---|---|---|---|---|---|---|---|---|
| | High-confidence level (6) | | All levels (4,5,6) | | High-confidence level (1) | | All levels (1,2,3) | |
| | Exp1 | Exp2 | Exp1 | Exp2 | Exp1 | Exp2 | Exp1 | Exp2 |
| P (R \|old item) | .36 | .30 | .48 | .41 | .03 | .05 | .08 | .11 |
| P (R \|similar item) | .12 | .09 | .22 | .17 | .23 | .25 | .30 | .32 |
| P (R \|new item) | .03 | .02 | .09 | .10 | .11 | .06 | .18 | .13 |
| Threshold correction | .24 | .22 | .25 | .25 | .20 | .21 | .22 | .22 |

**Table 4. Remember responses (with means and standard errors) and intercepts for Experiment 3 (With-New blocks).**

| | "Old" | | | | "New" | | | |
|---|---|---|---|---|---|---|---|---|
| | High-confidence level (6) | | All levels (4,5,6) | | High-confidence level (1) | | All levels (1,2,3) | |
| | M | SE | M | SE | M | SE | M | SE |
| P (R \|old item) | .53 | .06 | .59 | .05 | .03 | .01 | .10 | .02 |
| P (R \|similar item) | .13 | .03 | .18 | .04 | .44 | .07 | .51 | .07 |
| P (R \|new item) | .04 | .02 | .09 | .03 | .15 | .04 | .27 | .06 |
| Threshold correction | .40[a] | .06 | .41[a] | .06 | .41[b] | .07 | .42[b] | .07 |

| | M | SE | M | SE |
|---|---|---|---|---|
| ROC estimates | .47[c] | .07 | .46[d] | .08 |

Notes: 1. Threshold corrections for (a) and (b) are performed as follows:
a. P (R | old item called "old" ) – P (R | similar item called "old")
b. P (R | similar item called "new") – P (R | old item called "new")
2. The ROC-based estimates of y-intercept (c) and upper x-intercept (d)
3. *P (R | old item)*: Probability of making a "remember" response to an "old" item

**Table 5. Remember responses (with means and standard errors) and intercepts for Experiment 3 (No-New blocks).**

| | "Old" High-confidence level (6) | | "Old" All levels (4,5,6) | | "New" High-confidence level (1) | | "New" All levels (1,2,3) | |
|---|---|---|---|---|---|---|---|---|
| | M | SE | M | SE | M | SE | M | SE |
| P (R \|old item) | .50 | .05 | .56 | .05 | .07 | .01 | .12 | .02 |
| P (R \|similar item) | .12 | .03 | .15 | .04 | .43 | .06 | .50 | .06 |
| Threshold correction | .38[a] | .06 | .41[a] | .06 | .35[b] | .06 | .38[b] | .06 |

| | "Old" M | "Old" SE | "New" M | "New" SE |
|---|---|---|---|---|
| ROC estimates | .48[c] | .06 | .35[d] | .06 |

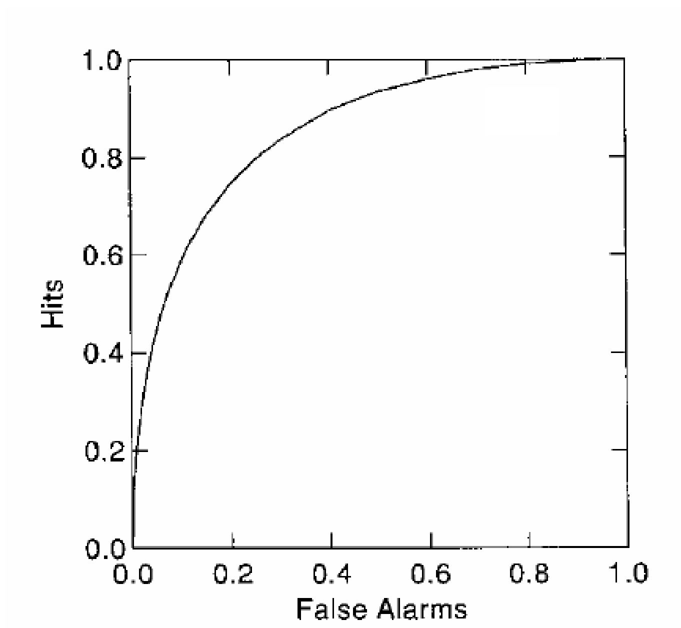
Notes: 1. Threshold corrections for (a) and (b) are performed as follows:

a. P (R | old item called "old" ) – P (R | similar item called "old")
b. P (R | similar item called "new") – P (R | old item called "new")

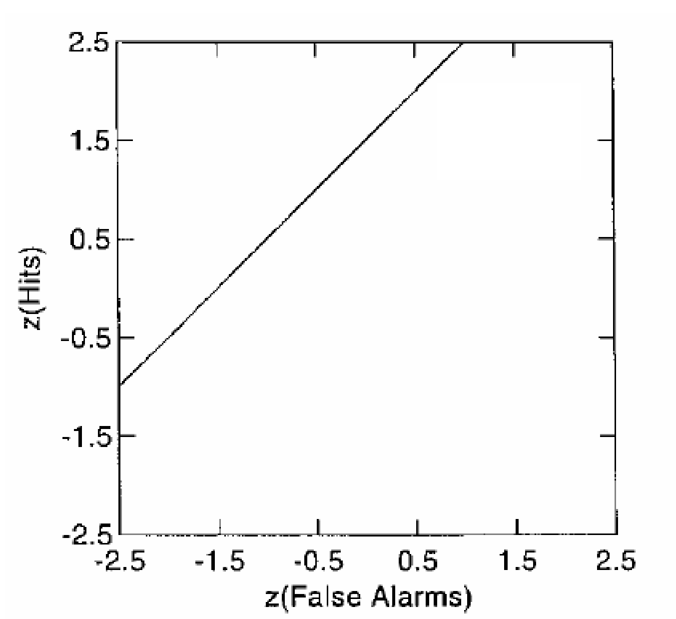
2. The ROC-based estimates of y-intercept (c) and upper x-intercept (d)
3. *P (R | old item)*: Probability of making a "remember" response to an "old" item

**Figure 1a. Hypothetical ROC (A) and z-ROC (B) curves predicted by a continuous familiarity model with equal-variance.**

**A.**



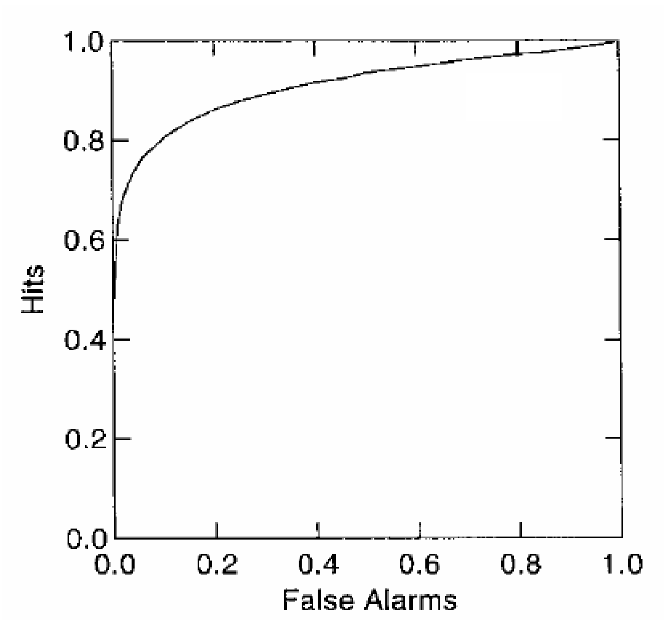

**B.**



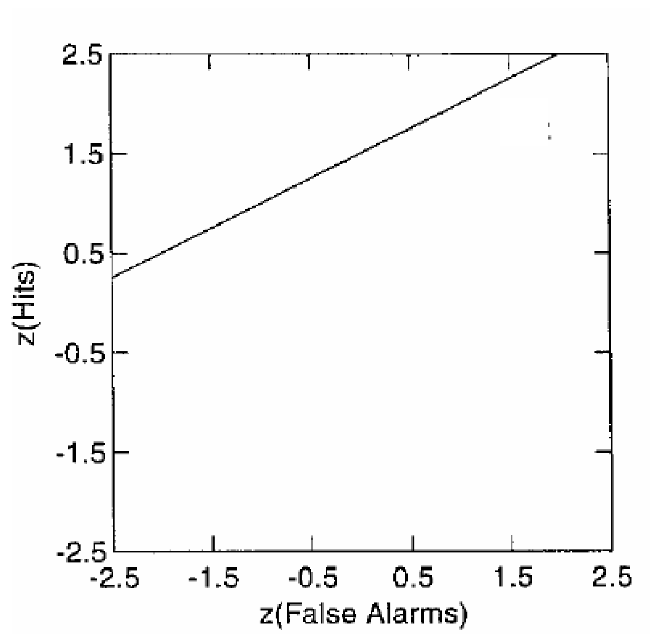

Note: The figures are adopted from Rotello, Macmillan, and Van Tassel, 2000.

**Figure 1b. Hypothetical ROC (A) and z-ROC (B) curves predicted by a continuous familiarity model with unequal-variance.**

**A.**




**B.**

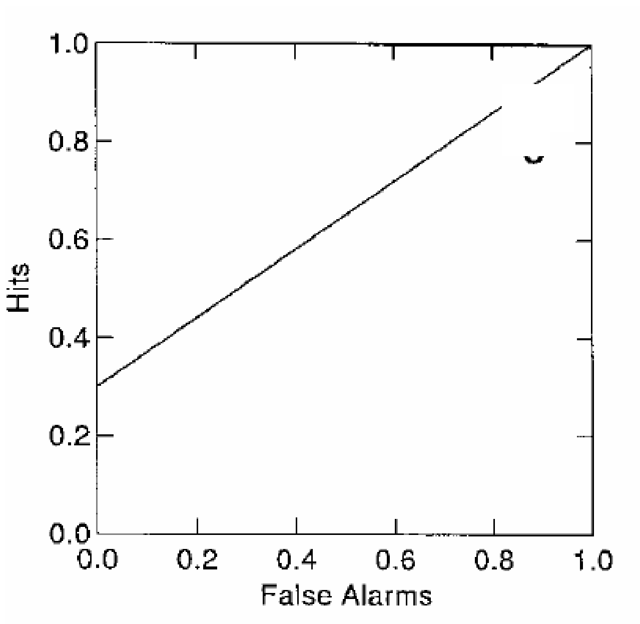



---

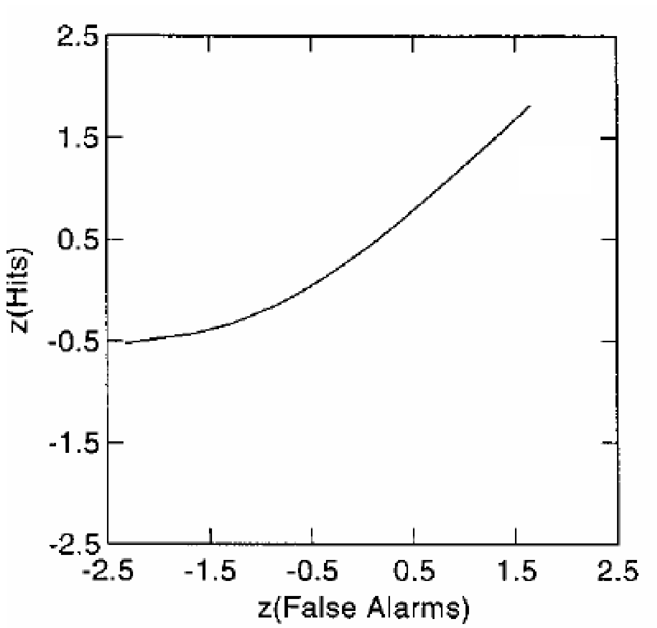
Note: The figures are adopted from Rotello et al., 2000

**Figure 2. Hypothetical ROC (A) and z-ROC (B) curves predicted by a high-threshold model.**

**A.**



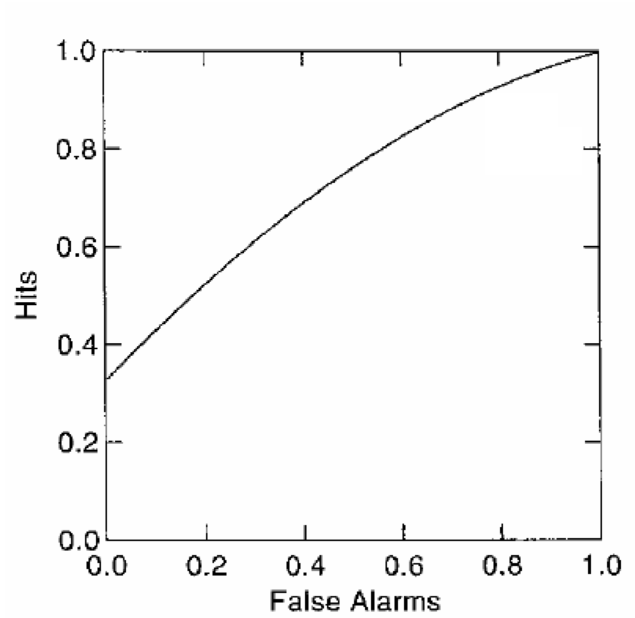

**B.**



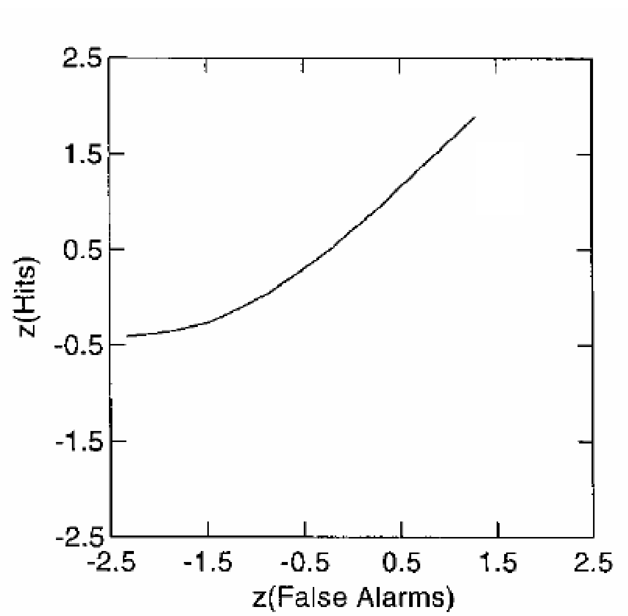

Note: The figures are adopted from Rotello et al., 2000.

**Figure 3. Hypothetical ROC and z-ROC curves predicted by a dual-process model.**

**A.**



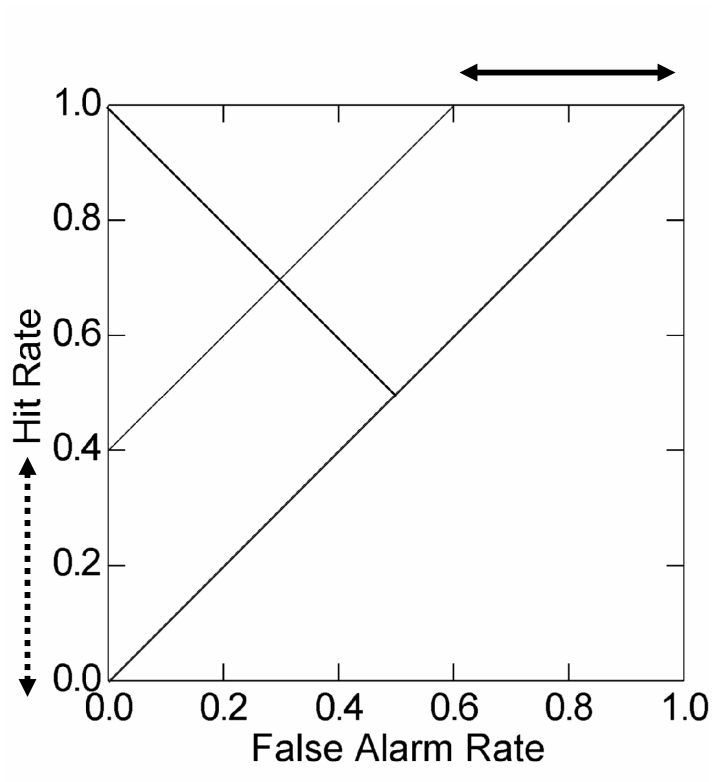

**B.**




---

Note: The figures are adopted from Rotello et al., 2000.

**Figure 4. Hypothetical ROC curve showing the use of recall-to-accept and recall-to-reject processes.**




---

Note: dashed line (y-intercept) shows the recall-to-accept processing and solid line (upper x-intercept) shows the recall-to-reject processing.

**Figure 5. Signal-detection model for the remember-know paradigm.**

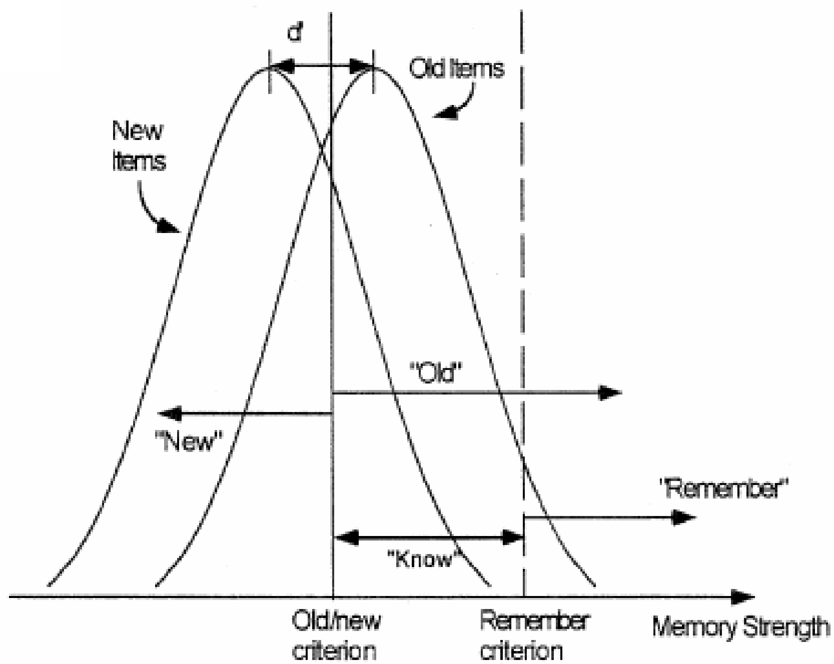


**Figure 6. The sum-difference theory of remembering and knowing (STREAK).**

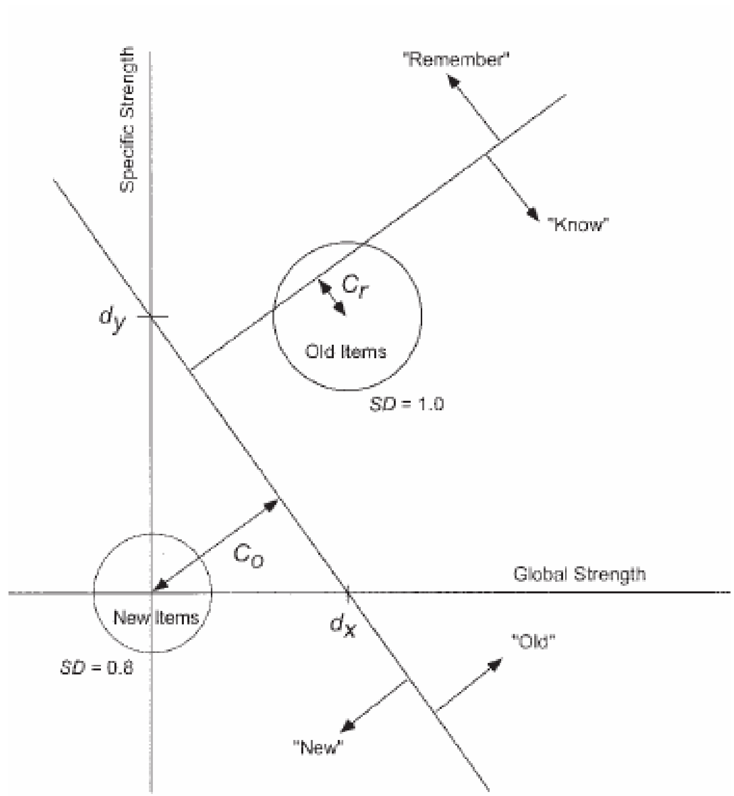
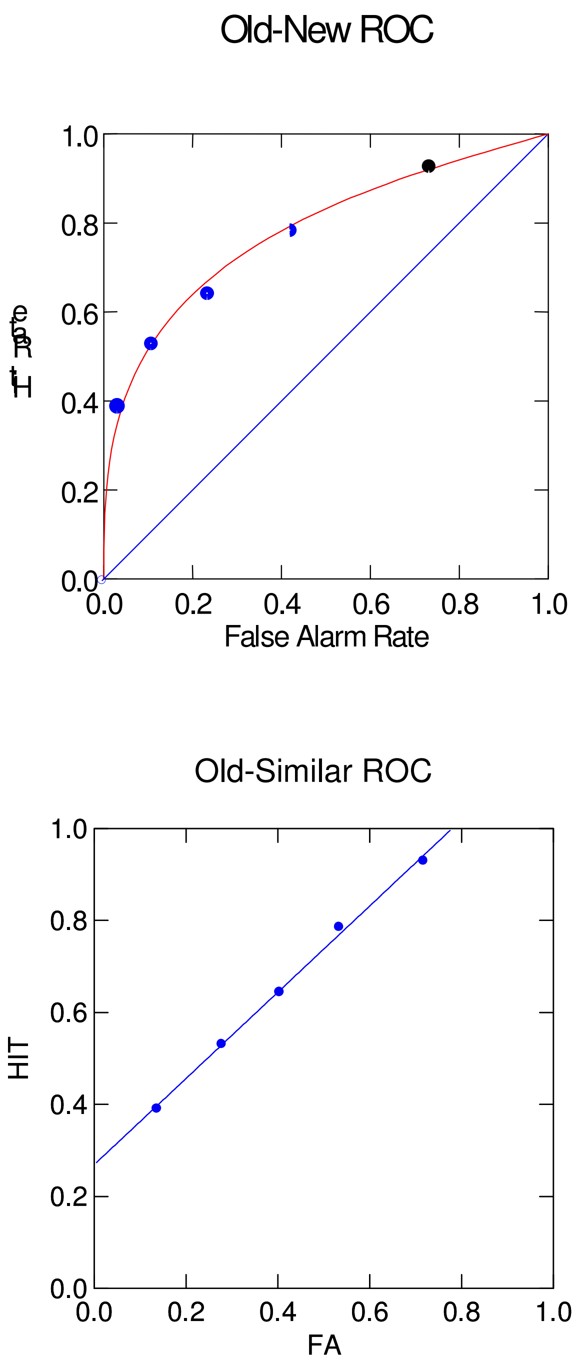


**Figure 7. ROC curves for the old-new and the old-similar data in Experiment 1.**




Note: *Old-new ROC*: Hit rate to old items is plotted against false alarm rate to new items.
*Old-similar ROC*: Hit rate to old items is plotted against false alarm rate to similar items.

**Figure 8. z-ROC curves for the old-new and the old-similar data in Experiment 1.**

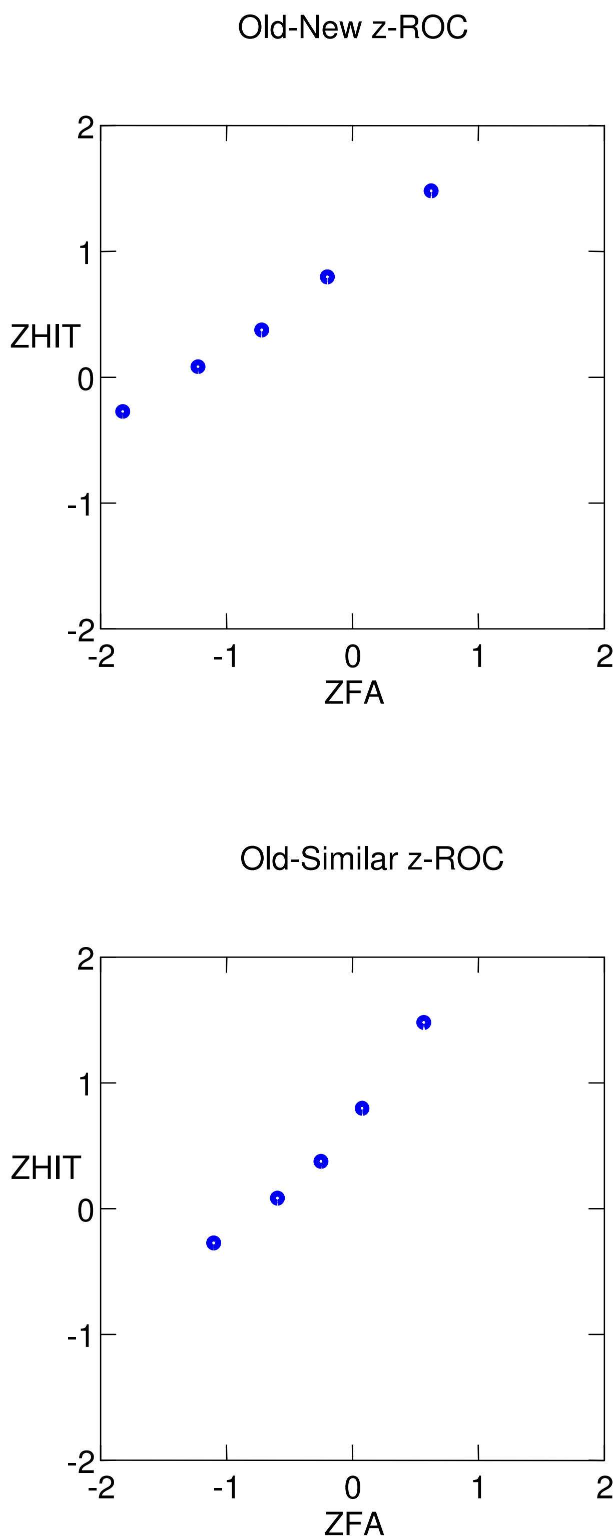


**Figure 9. ROC curves for the old-new and the old-similar data in Experiment 2.**

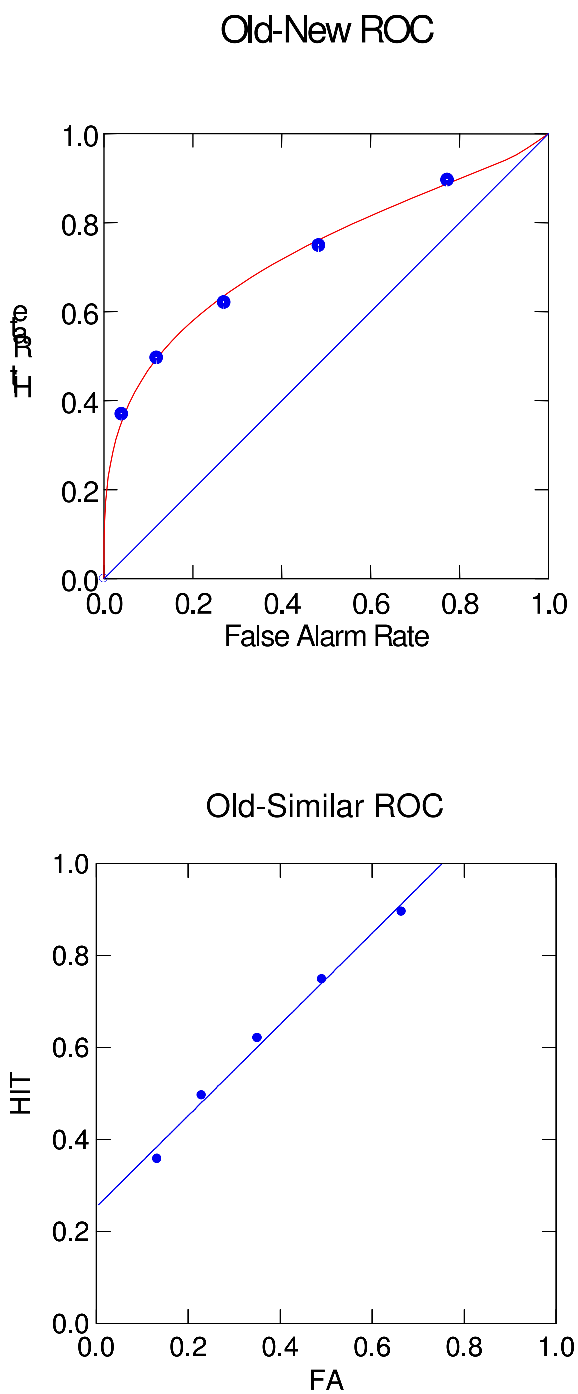


**Figure 10. z-ROC curves for the old-new and the old-similar data in Experiment 2.**

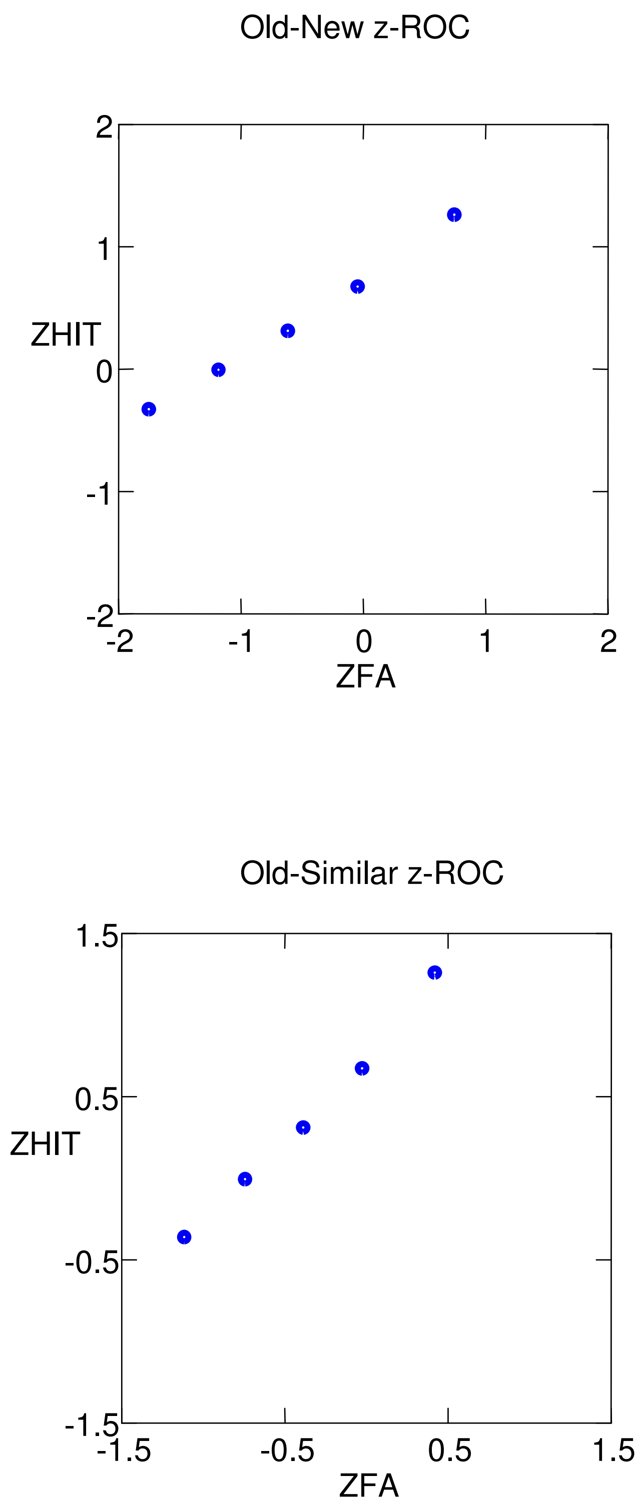


**Figure 11. The old-similar ROC curves from Heathcote et al. (2006)**

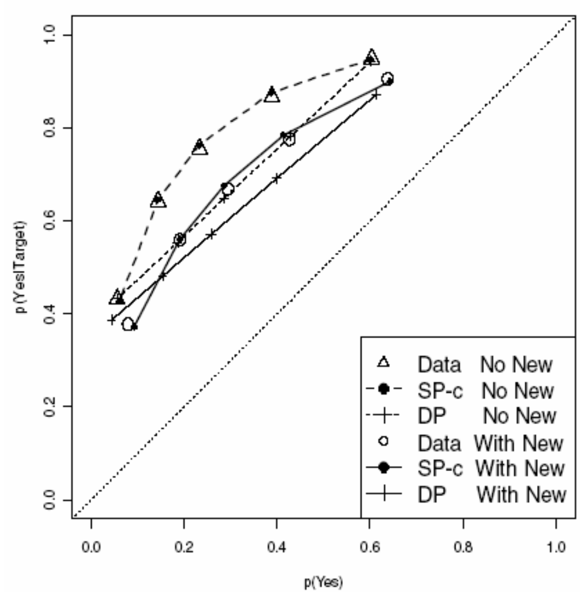
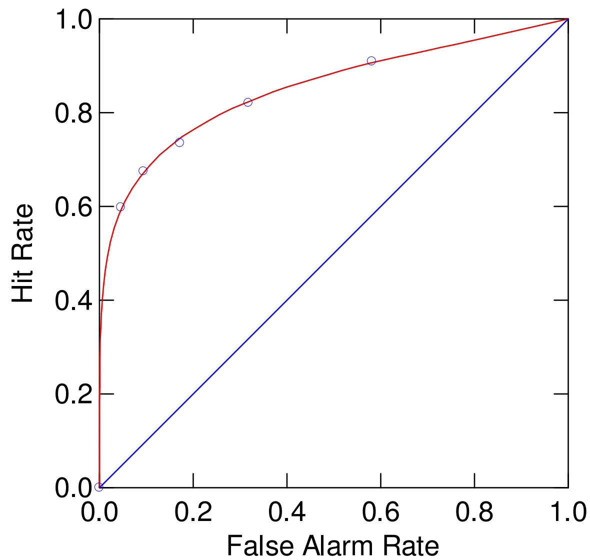


**Figure 12. ROC and z-ROC curves for the old-new data in Experiment 3.**

Old-New ROC




Old-New z-ROC

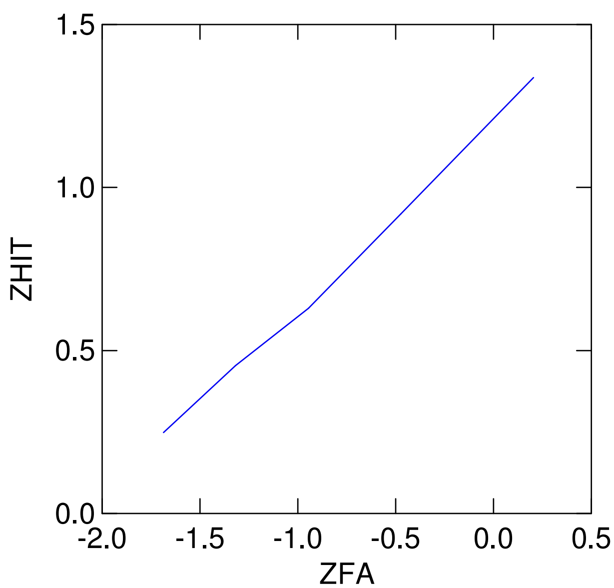
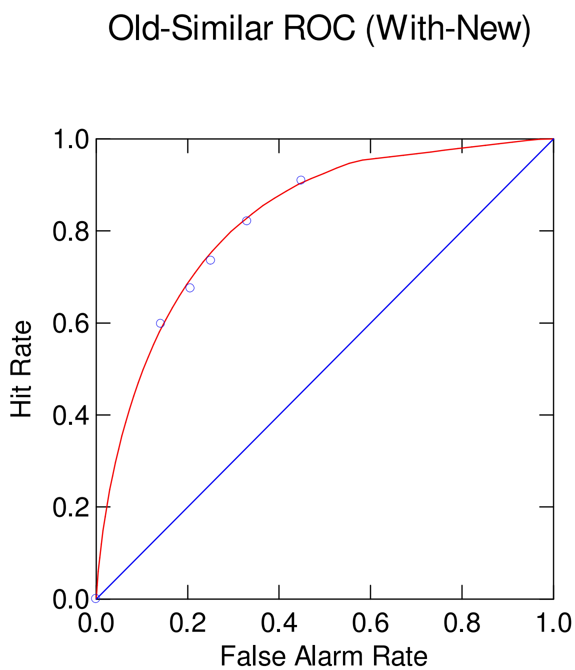


**Figure 13. ROC curves for the old-similar data for With-New and No-New blocks in Experiment 3.**



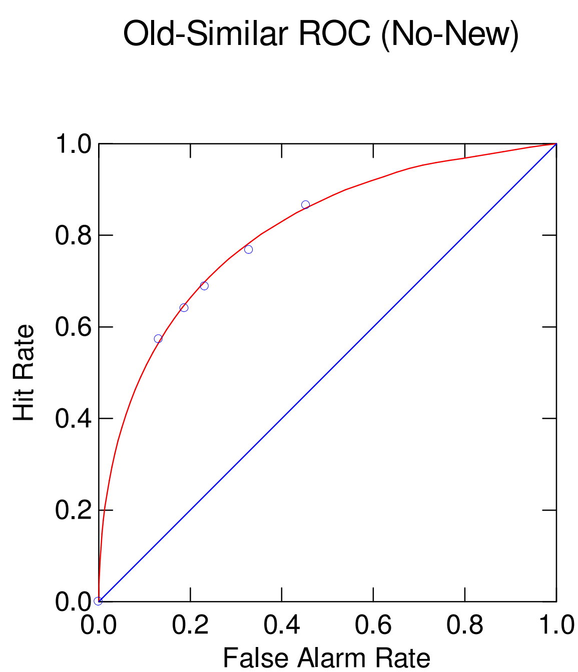
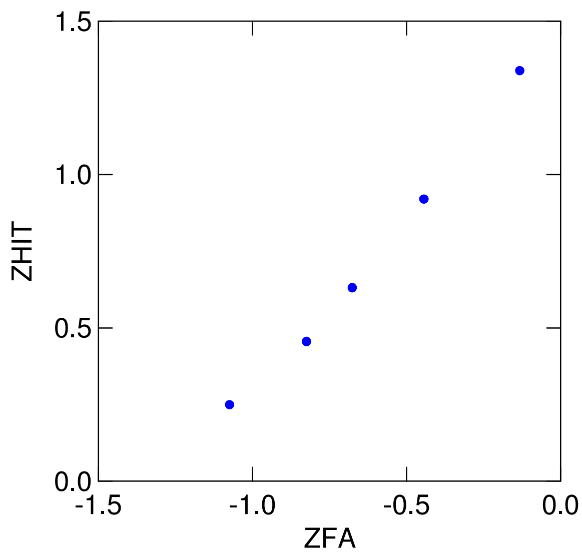



**Figure 14. z-ROC curves for the old-similar data for With-New and No-New blocks in Experiment 3.**

Old-New z-ROC (With-New)




Old-Similar z-ROC (No-New)

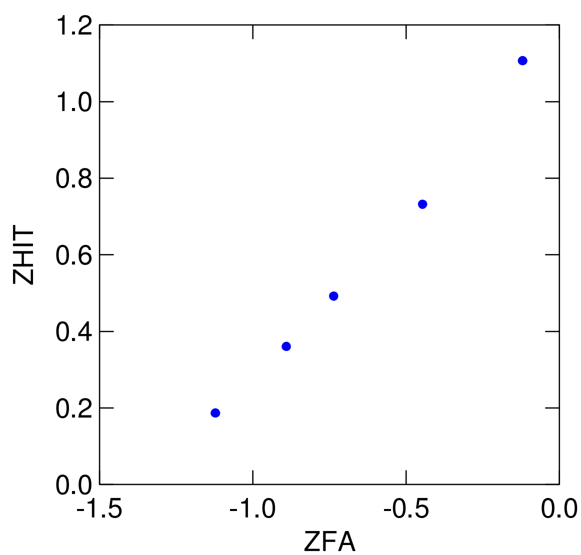
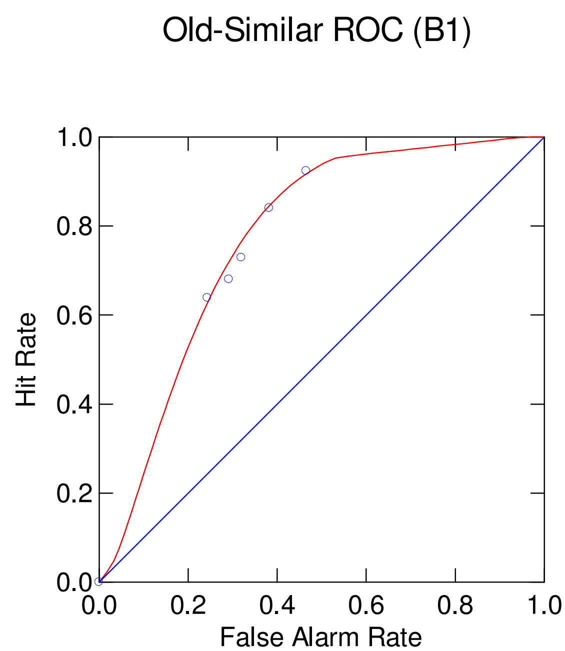
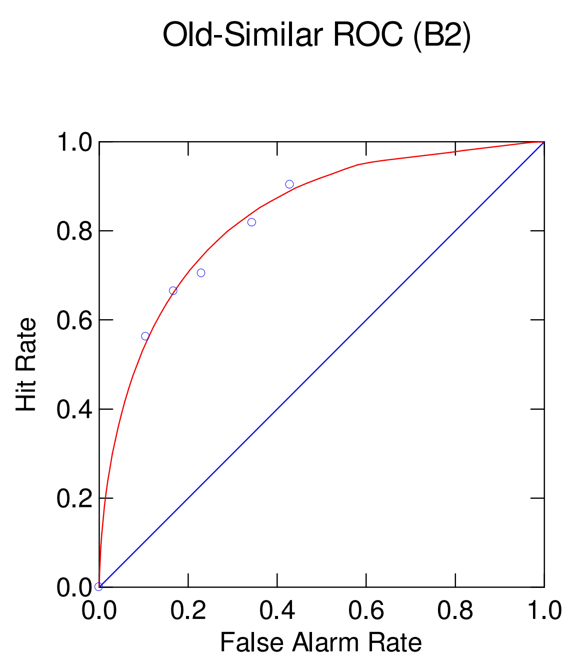


**Figure 15. ROC curves for the old-similar data for With-New first (B1) and No-New first blocks (B2) in Experiment 3.**



## REFERENCES


Benjamin, A. S. (2005). Recognition memory and introspective remember/know judgments: Evidence for the influence of distractor plausibility in "remembering" and a caution about purportedly nonparametric measures. *Memory and Cognition, 33*(2), 261-269.

Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory and Cognition*, *24*(4), 523-533.

Dunn, J. C. (2004). Remember-know: A matter of confidence. *Psychological Review*, *111*(2), 524-542.

Gardiner, J. M., & Richardson-Klavehn, A. (2000). Remembering and knowing. In E. E. Tulving, E. Fergus, & I. M. Craik (Eds.), *The Oxford handbook of memory* (pp.229-244). New York: Oxford Univ. Press.

Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (2002). Recognition memory and decision processes: A meta-analysis of remember, know and guess responses. *Memory*, *10*(2), 83-98.

Ghetti, S. (2003). Memory for nonoccurrences: The role of metacognition. *Journal of Memory and Language, 48*(4), 722-739.

Glanzer, M., Kim, K., & Hilford, A. (1999). Slope of the receiver-operating characteristic in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(2), 500-513.

Heathcote, A., Raymond, F., & Dunn, J. (2006). Discrete recollection in recognition memory: Evidence from receiver operating characteristics. *Journal of Memory and Language, 55,* 495-514.

Hirshman, E., & Henzler, A. (1998). The role of decision processes in conscious recollection. *Psychological Science, 9*(1), 61-65.

Hirshman, E., & Master, S. (1997). Modeling the conscious correlates of recognition memory: Reflections on the remember-know paradigm. *Memory and Cognition*, *25*(3), 345-351.

Jones, T. C. (2005). Study repetition and the rejection of conjunction lures. *Memory, 13*(5), 499-515.

Lampinen, J. M., Odegard, T. N., & Neuschatz, J. S. (2004). Robust recollection rejection in the memory conjunction paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(2), 332-342.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide.* New York: Cambridge Univ. Press.

Macmillan, N. A., Rotello, C. M., & Verde, M. F. (2005). On the importance of models in interpreting remember-know experiments: Comments on Gardiner et al.'s (2002) meta-analysis. *Memory*, *13*(6), 607-621.

Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory and Cognition*, *21*(1), 89-102.

Ratcliff, R., McKoon, G., & Tindall, M. (1994). Empirical generality of data from recognition memory receiver-operating characteristic functions and implications for the global memory models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(4), 763-785.

Rotello, C. M., & Heit, E. (1999). Two process models of recognition memory: Evidence for recall-to-reject? *Journal of Memory and Language, 40*, 432-453.

Rotello, C. M. (1999). Metacognition and memory for nonoccurrence. *Memory, 7*(1), 43-63.

Rotello, C. M., Macmillan, N. A., & Reeder, J. A. (2004). Sum-difference theory of remembering and knowing: A two-dimensional signal-detection model. *Psychological Review, 111*(3), 588-616.

Rotello, C. M., Macmillan, N. A., & Van Tassel, G. (2000). Recall-to-reject: Evidence from ROC curves. *Journal of Memory and Language, 43*, 67-88.

Rotello, C. M., Macmillan, N. A., Reeder, J. A., & Wong, M. (2005). The remember response: Subject to bias, graded, and not a process-pure indicator of recollection. *Psychonomic Bulletin and Review, 12*(5), 865-873.

Rotello, C. M. (2001). Recall processes in recognition memory. In Medin, D. L. (Eds.) *The psychology of learning and motivation: Advances in research and theory* (pp. 183-221). San Diego: Academic Press.

Strack, F., & Bless, H. (1994). Memory for nonoccurrences: Metacognitive and presuppositional strategies. *Journal of Memory and Language, 33*, 203-217.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology, 26,* 1-12.

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*(6), 1341-1354.

Yonelinas, A. P., Dobbins, I., Szymanski, M. D., Dhaliwal, H. S., & King, L. (1996). Signal detection, threshold and dual-process models of recognition memory: ROCs and conscious recollection. *Consciousness and Cognition, 5,* 418-441.

Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory and Cogniton, 25*(6), 747-763.

Yonelinas, A. P. (2001). Consciousness, control and confidence: The 3 Cs of recognition memory. *Journal of Experimental Psychology: General, 130*(3), 361-379.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46,* 441-517.

Wixted, J. T., Stretch, V. (2004). In defense of the signal-detection interpretation of remember/know judgments. *Psychonomic Bulletin and Review, 11*(4), 616-641.

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review, 114 (1),* 152-176.