CONSCIOUSNESS AND COGNITION **5,** 418–441 (1996)
ARTICLE NO. 0026

Signal-Detection, Threshold, and Dual-Process Models of Recognition Memory: ROCs and Conscious Recollection

ANDREW P. YONELINAS,[1] IAN DOBBINS, MICHAEL D. SZYMANSKI, HARPREET S. DHALIWAL, AND LING KING

Department of Psychology, University of California at Davis, Davis, California 95616

Threshold- and signal-detection-based models have dominated theorizing about recognition memory. Building upon these theoretical frameworks, we have argued for a dual-process model in which conscious recollection (a threshold process) and familiarity (a signal-detection process) contribute to memory performance. In the current paper we assessed several memory models by examining the effects of levels of processing and the number of presentations on recognition memory receiver operating characteristics (ROCs). In general, when the ROCs were plotted in probability space they exhibited an inverted U shape; however, when they were plotted in *z* space they exhibited a U shape. An examination of the ROCs showed that the dual-process model could account for the observed ROCs, but that models based solely on either threshold or signal-detection processes failed to provide a sufficient account of the data. Furthermore, an examination of subjects' introspective reports using the remember/know procedure showed that subjects were aware of recollection and familiarity and were able to consistently report on their occurrence. The remember/know data were used to accurately predict the shapes of the ROCs, and estimates of recollection and familiarity derived from the ROC data mirrored the subjective reports of these processes.

Threshold- and signal-detection-based theories represent two fundamentally different ways of conceptualizing human memory. With threshold theory, it is assumed that there is some probability that previously studied items will exceed a memory threshold. If an item exceeds the threshold then it is in a discrete memory state. If an item does not exceed the threshold then it is not remembered, but it may still be endorsed as old on the basis of a random guess. Although there are many ways a threshold process may be realized, one plausible way is to assume that memory reflects a discrete retrieval process that provides qualitative information about a previous event. Thus, if we recollect what a person said or what they were wearing, we may be confident that we have met them before. If we cannot retrieve anything about the person then memory has failed, and we either respond that we do not recognize the person or we simply make a random guess.

A very different way of thinking about memory is to assume that there is no discrete memory threshold or state, but rather that memory reflects a type of educated guessing. For example, signal-detection-based models assume that we can place items on a familiarity continuum such that studied items fall on the high end of the continuum and new items fall on the low end of the continuum. By such a model, there is no threshold above which we can be certain that an item was studied, because the famil-

[1] To whom correspondence should be addressed at Department of Psychology, University of California, Davis, CA 95616. E-mail: apyonelinas@ucdavis.edu.

1053-8100/96 $18.00

iarity distributions of old and new items overlap. Memory judgments are made by setting some level of familiarity as a response criterion and accepting items that exceed this criterion as having been studied. As with threshold theory, there are many ways in which a signal-detection process could be realized. However, most current recognition memory models (e.g., connectionist models and global models such as TODAM (Murdock, 1982), SAM (Gillund & Shiffrin, 1984), and MINERVA2 (Hintzman, 1986)) assume that a comparison of each test item to long-term memory produces a continuous familiarity value that is used to make recognition judgments, and this process is well described by signal-detection theory. The notion underlying these models is that studied items will be more similar to what has been stored in memory than will new items; thus, old items will be the most familiar.

Another possibility is that recognition memory judgments rely on a threshold process *and* a signal-detection process. That is, items can be recognized either if they are recollected or if they are sufficiently familiar. The dual-process notion of memory, which dates back to Aristotle, has been elaborated by several contemporary cognitive psychologists (e.g., Atkinson & Juola, 1974; Jacoby & Dallas, 1981; Mandler, 1980) and has recently enjoyed renewed interest in the light of neuropsychological findings that amnesics suffer from a deficit in recollection that leaves familiarity largely intact (e.g., Huppert & Piercy, 1976; Verfaellie & Treadwell, 1993).

The dual-process model that we explore in the current study assumes that recollection is well described as a threshold process, and familiarity is well described as a signal-detection process (Yonelinas, 1994). The notion is that when a subject consciously recollects information about a prior event they are in a discrete memory state (they either retrieve qualitative information about the event or they do not). Subjects may recollect several different aspects of a study event, but recollection of any aspect of the event is expected to lead to a confident recognition judgment. In contrast, memory judgments can also be made on the basis of how familiar an item seems in the experimental context. Familiarity is assumed to be continuous in nature such that the more familiar an item seems, the more confident the subject should be that it was studied. Although we will focus on recognition memory in the current study, familiarity may also support performance on other tests of memory. For example, familiarity may lead a name to seem famous (Jacoby, Woloshyn, Kelley, 1989), a shape to be preferred (Kunst-Wilson & Zajonc, 1980), or a word to come to mind readily in a fragment completion (Warrington & Weiskrantz, 1974) or a perceptual identification task (Jacoby & Dallas, 1981).

We begin by describing the threshold, signal-detection and dual-process models in more detail and then present the results from two recognition memory experiments designed to test these models. In Experiments 1 and 2, recognition performance was plotted as a function of response confidence to determine how well the different models accounted for recognition performance. In addition, in Experiment 2, subjective reports of recollection and familiarity were examined to determine if the processes underlying the dual-process model were available to subjective awareness.

THRESHOLD MODELS

There are three different threshold models that we will consider: the high-threshold, 2-threshold, and 2-equal-threshold models. Although there are other threshold models (see Swets, 1986), these three are the simplest and most frequently used. The high-threshold and the 2-threshold models underlie the multinomial models (Batchelder & Riefer, 1990) that have been popular in studies of recognition and source memory (e.g., Johnson, Kounios, & Reeder, 1994; Bayen, Murnane, & Erdfelder, 1996).

With the high-threshold model, a studied item will be accepted if it exceeds the memory threshold or on the basis of a guess. As previously described, a person may recollect an item, but in the absence of recollection, they may simply guess. The probability that a studied item will be accepted is equal to the probability that it is remembered (R) plus the probability that when it is not remembered ($1 - R$), it is accepted on the basis of a guess (G):

$$P(\text{``old''}|\text{old}) = R + (1 - R)\,G.$$

Items that have not been studied (i.e., new items) do not exceed the memory threshold, but they may be accepted on the basis of a guess. Thus, the probability of a false alarm can be written

$$P(\text{``old''}|\text{new}) = G.$$

The 2-threshold model is similar to the high-threshold model except that subjects are assumed to be able to recognize that new items were not in the study list. For example, imagine coming across your own name in a recognition memory test list. If you do not remember encountering your name in the study list, you may reason that it probably was not studied. Thus, a new item will only be incorrectly accepted as old if it is not recollected as new ($1 - R_n$). If the probability of recognizing old and new items is equal to R_o and R_n, respectively, then

$$P(\text{``old''}|\text{old}) = R_o + (1 - R_o)\,G$$

$$P(\text{``old''}|\text{new}) = G(1 - R_n).$$

Because there are two different recognition terms, the 2-threshold model requires one more free parameter than the high-threshold model. However, the model can be simplified by assuming that the probability of recognizing old and new items is equal (i.e., $R_o = R_n$). We will refer to this as the 2-equal-threshold model. This model may be represented in the following way:

$$P(\text{``old''}|\text{new}) = R + (1 - R)\,G$$

$$P(\text{``old''}|\text{new}) = G\,(1 - R).$$

SIGNAL-DETECTION MODELS

In contrast to the threshold models which are based on linear equations, signal-detection models (e.g., Banks, 1970; Green & Swets, 1966; Murdock, 1965; Swets, Tanner, & Birdsall, 1961; Wickelgren & Norman, 1966) assume that performance is described by nonlinear equations. Although there are numerous potential nonlinear models that could be described as signal-detection models, we will restrict our discussion to the most common class which are based on normal distributions. We will begin by describing the equal-variance signal-detection model and then describe the

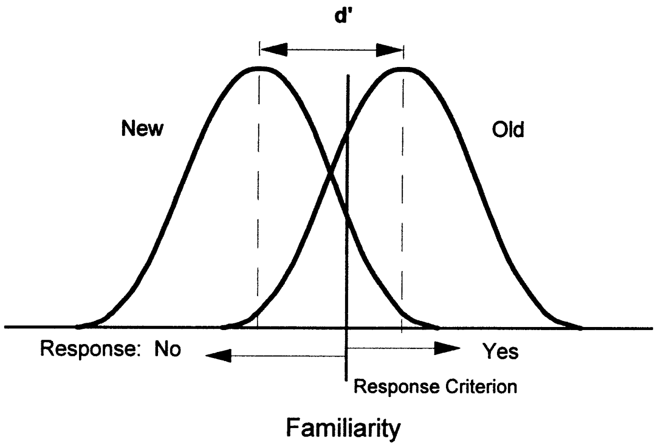

FIG. 1. Familiarty distributions for the equal-variance signal-detection model.

slightly more complex unequal-variance signal-detection model. As previously mentioned, the notion underlying signal-detection theory is that items can be represented as falling on a familiarity continuum. Presumably, new items fall on the lower end of the familiarity continuum, and studied items fall on the higher end of the continuum. However, there is some variability from one item to the next such that the familiarities of old and new items are normally distributed and overlap each other, as in Fig. 1. To discriminate between old and new items, subjects are assumed to select some level of familiarity (i.e., the response criterion) so that only the items exceeding this level are accepted as old.

Signal-detection theory assumes that the hit rate is equal to the proportion of the old item distribution that exceeds the response criterion. That is, subjects will correctly respond ''old'' to an old item if its familiarity value exceeds the criterion

$$P(\text{``old''}|\text{old}) = (F_o > c).$$

Similarly, the false alarm rate is equal to the proportion of the new item distribution that exceeds the response criterion

$$P(\text{``old''}|\text{new}) = (F_n > c).$$

The model illustrated in Fig. 1 is referred to as an equal-variance signal-detection model because the variance of the old and new item distributions are equal. Given that the distributions are of equal variance, $F_o > c$ and $F_n > c$ are determined by d' (the distance between the means of the new and old distributions) and c (the response criterion). The equal-variance assumption implies that studying an item leads to a constant increase in familiarity.

However, one can relax the equal-variance assumption and allow the variance of the old item distribution to differ from that of the new item distribution. This unequal-variance signal-detection model can be described by assuming that the new item distribution has a variance of 1 and letting the variance of the old item distribution be equal to V_o. If the variance distributions are unequal then performance will be determined by d', c, and V_o. For a more complete description of signal detection theory see Green and Swets (1966) or Macmillan and Creelman (1991).

A DUAL-PROCESS MODEL

An alternative approach is to assume that recognition judgments can be based on a threshold process (e.g., the recollection of qualitative information about the study event) and a signal-detection process (e.g., the assessment of familiarity). For simplicity, the familiarity distributions are assumed to be normal and of equal variance, and recollection is assumed to lead to high confidence recognition responses. Moreover, in standard tests of recognition memory, recollection is assumed to reflect a high-threshold model of the type described earlier. Note, however, that alternative threshold assumptions may be more appropriate for different memory tasks (see Yonelinas, in press). The two processes are assumed to serve as independent bases for recognition judgments such that the probability of recognizing an old item can be written

$$P(\text{``old''}|\text{old}) = R + (1 - \text{R})\,(F_o > c).$$

Given that the subject adopts a particular response criterion, they will also accept a certain proportion of the new items. The probability of incorrectly accepting a new item is equal to the probability that its familiarity exceeds the response criterion:

$$P(\text{``old''}|\text{new}) = (F_n > c).$$

If the familiarity distributions are normal and of equal variance, as in the equal-variance signal-detection model, then performance will be a function of R, d', and c. The model is essentially a hybrid of the threshold and signal-detection models previously discussed and can be thought of as a high-threshold model in which the random guessing process is replaced by an educated guessing process (i.e., the assessment of familiarity).

USING ROCS TO ASSESS THE MODELS

The most direct way of testing these models is to examine empirical receiver operating characteristics (ROCs). An ROC is the function that relates the hit rate to the false alarm rate—where the hit rate is plotted on one axis and the false alarm rate on the other. Typically, subjects are required to rate the confidence of their recognition judgments, and performance is plotted as a function of response confidence. For example, the first point on the function is determined by adopting a very strict response criterion—including only the most confidently remembered items as hits and false alarms. The second point reflects a slightly more relaxed response criterion—the most confidently remembered items as well as the second most confidently remembered items would be treated as hits and false alarms. By examining performance at a number of different levels of response confidence we can plot the ROC and examine the relationship between hits and false alarms.

The ROC is extremely useful in assessing the alternative models of recognition, because the fundamental difference between the memory models is the assumed relationship between hits and false alarms. Because the models posit different equations

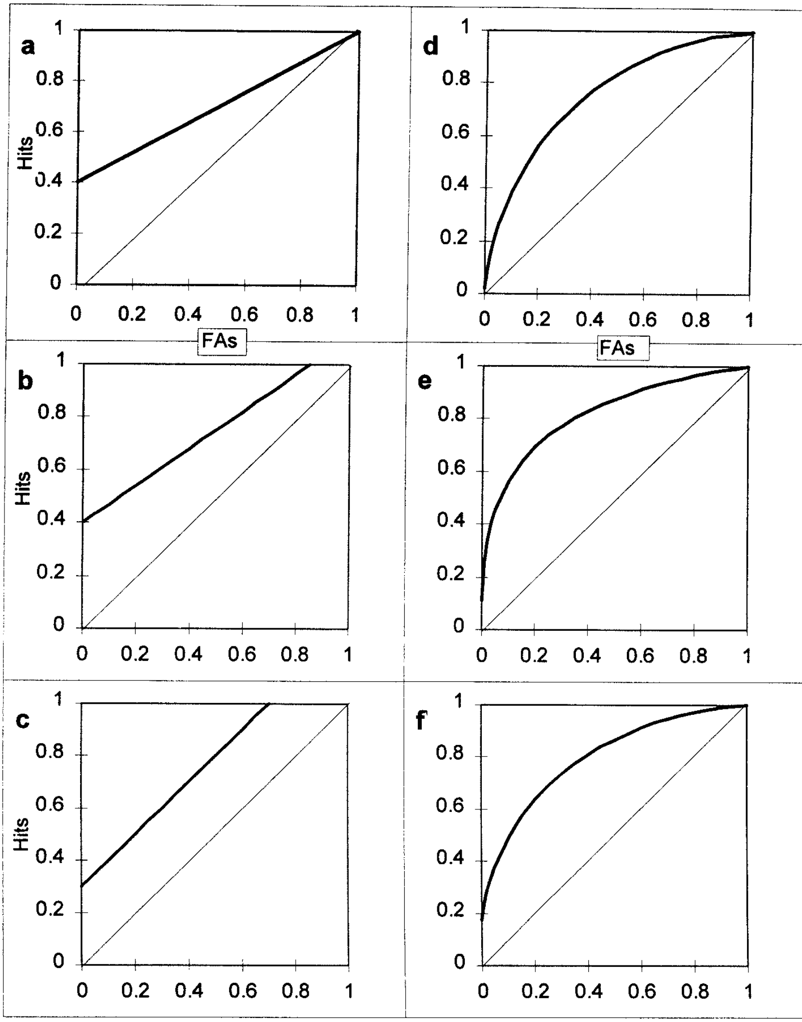

FIG. 2. Predicted ROCs for the (a) high-threshold, (b) 2-threshold, (c) 2-equal-threshold, (d) equal-variance signal-detection, (e) unequal-variance signal-detection, and (f) dual-process models.

for hits and false alarms, they predict very different ROCs. Thus, by examining ROCs we can determine which model provides the best account of recognition performance.

Theoretical ROCs that were generated using the threshold models, the signal-detection models, and the dual-process model are presented in Fig. 2. These functions reflect predicted performance at a constant level of memory sensitivity, as response bias or guessing is varied. Examination of Fig. 2a shows that the high-threshold model predicts a linear ROC that approaches the 1,1 intercept. The function's *y* intercept reflects the probability that an old item is recollected, and this value can vary from 0 to 1.0. The 2-threshold model (Fig. 2b) also predicts a linear function, but the *y* and upper-*x* intercepts (R_o and R_n, respectively) are free to vary between 0 and 1.0. The 2-equal-threshold model (Fig. 2c) predicts a liner ROC, but because R_0 is equal to R_n it predicts a function that is parallel to the diagonal.

The equal-variance signal-detection model (Fig. 2d) predicts a curvilinear ROC that is symmetrical along the diagonal. A decrease in memory performance (d') would lead to a function that is closer to the diagonal, but the function will always be symmetrical. In contrast, the unequal-variance signal-detection model predicts a curvilinear ROC that is not symmetrical along the diagonal. If the variance of the old item distribution is greater than that of the new item distribution then the function will be pulled toward the bottom left corner, as in Fig. 2e. The degree of asymmetry will increase as the variance of the old item distribution increases. If the variance of the old item distribution becomes less than that of the new item distribution then the function will be asymmetrical in the opposite direction and will be pulled toward the upper right corner.

The dual-process model (Fig. 2f) predicts an ROC that it is curved and asymmetrical along the diagonal. The function can be generated using the dual-process equations presented earlier or by adding the independent contribution of recollection to the hit rate of the equal-variance signal-detection ROC. Thus, each hit rate from the signal-detection ROC is increased by ($R - R*$ (the hit rate)). A comparison of Figs. 2e and 2f shows that the model can produce an ROC that looks very much like that of the unequal-variance signal-detection model. However, there are subtle differences between the ROCs predicted by the two models, and there are cases in which the models predict very different functions. Because the dual-process model assumes that there is a threshold process that contributes to performance, it predicts an ROC that is slightly more linear than that predicted by the signal-detection models. This linearity becomes most noticeable when recollection increases and familiarity decreases. In fact, if there is no contribution of familiarity to memory performance, the model predicts a function that looks like that of the high-threshold model (i.e., a straight line).

EXPERIMENT 1

In Experiment 1 we examined recognition memory ROCs to determine which of the threshold, signal-detection, and dual-process models provided the most accurate account of memory performance. Because the dual-process and the unequal-variance signal-detection models can produce very similar ROCs we wanted to examine performance under conditions where we expected the two models to predict divergent functions. To do this, we examined the effects of levels of processing (Craik & Lockhart, 1972) by assessing performance for items studied under semantic and perceptual encoding conditions. Semantic processing, compared to perceptual processing, tends to increase the likelihood that subjects will recollect aspects of the study event (e.g., Gardiner, 1988; Toth, 1996). If recollection is a threshold process then the ROC for items encoded under semantic encoding conditions should become more linear than would be predicted by a signal-detection process alone. In contrast, because recollection is less likely in the perceptual encoding condition, the ROC should be fit reasonably well by both the signal-detection and the dual-process models.

Subjects in Experiment 1 heard two lists of words. For one list (perceptual encoding) subjects were instructed to count how many syllables were in each word. For the other list (semantic encoding) they were instructed to rate how pleasant the meaning of each word was. After the study phase, subjects were presented with a mixture of studied words and new words and asked to make recognition judgments on a 6-point confidence scale from 1 (sure it was new) to 6 (sure it was old). ROCs for items from the semantic and perceptual encoding conditions were plotted as a function of response confidence.

The models were assessed by examining the linearity of the observed ROCs and by fitting each model to the observed data. Performance was plotted on probability coordinates (ROCs) and *z* coordinates (z-ROCs). Because the threshold models are based on simple linear equations they predict linear ROCs. Z-ROCs were plotted because they provided a convenient way of assessing the signal-detection-based models. If performance is well described by signal-detection theory and the distributions of old and new items are normal (an assumption underlying both signal-detection models) then the z-ROCs should be straight lines.

If the threshold models are correct then the ROCs should be linear and the z-ROCs should exhibit a U shape. However, if the signal-detection-based models are correct then the ROCs should exhibit an inverted U shape and the z-ROCs should be linear. In contrast, if the dual process model is correct, and there is a signal-detection and a threshold process that contribute to performance, then the ROCs should exhibit a slight inverted U shape in probability space and a slight U shape in *z* space. Although the dual-process model and the unequal-variance signal-detection model predict very similar ROCs, the U shape in *z* space that is predicted by the dual-process model should become particularly noticeable in the semantic encoding condition, because recollection is expected to be greatest under these conditions.

In addition to assessing the qualitative predictions of the dual-process model, the ROCs were also used to assess the quantitative fit of the dual-process model and to contrast it with the threshold- and signal-detection-based models. The analysis was conducted by examining how well each model fit the observed data, with the goal of finding the simplest model that accounted for the most variance for each ROC.

Method

Subjects. Eighteen students participated in the experiment for credit in an undergraduate psychology course at the University of California, Davis. Subjects were treated in accordance with the university's human subjects review committee guidelines and could withdraw from participation at any time without penalty. Subjects were debriefed following completion of the experiment.

Materials. Two-hundred and forty words were selected from the Oxford Dictionary. Words were between one and four syllables in length and had word frequency counts between 3 and 15 per million (Kucera-Francis, 1967) and concreteness values (Coltheart, 1981) between 500 and 670. From this pool of items, 240 items were randomly selected and assigned to three lists of 80 words each.

Design and procedure. During the study phase, two lists of 80 words each were presented auditorially one after another using a cassette player. The words were read by a male voice at a rate of one item every 3.5 s. Subjects processed each list of

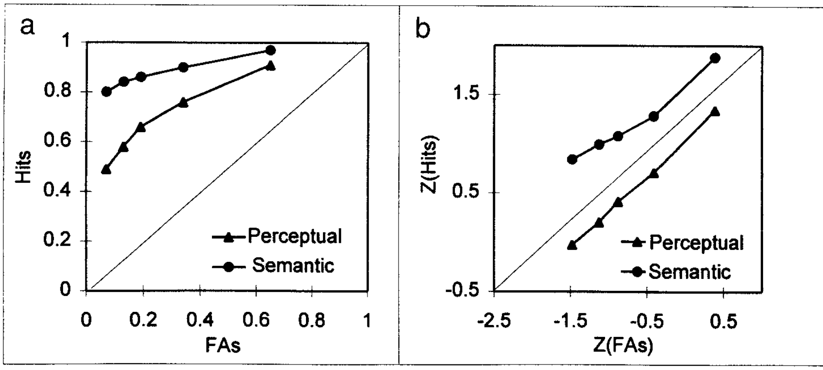

FIG. 3. Recognition memory ROCs from Experiment 1 for perceptually and semantically encoded items plotted on (a) probability coordinates and (b) *z* coordinates.

words either semantically or perceptually. For the semantic condition, subjects were instructed to judge how pleasant each word was using a 4-point scale, with 4 representing very pleasant words and 1 representing very unpleasant words. For the perceptual condition, subjects counted the number of syllables in each word. Subjects made their responses verbally in both encoding conditions. The order in which the semantic and perceptual tasks was performed was counterbalanced across subjects. Additionally, assignment of the three lists of words to the semantic, perceptual, and new conditions was counterbalanced such that each list served in each condition for an equal number of subjects.

Following the study phase, subjects were given a recognition memory test. All of the studied items and 80 new items were presented one at a time in a random order on an IBM-compatible computer in lowercase letters. The character size of the stimuli was approximately 5 $\times$ 5 mm, and the viewing distance was approximately .5 m. Subjects made recognition memory judgments on the computer keyboard using a 6-point confidence scale. Subjects were instructed to respond '6' if they were sure the item was presented in either of the study lists, '5' if they were less sure, and '4' if very unsure. They were instructed to respond '1' if they were sure the word was not presented earlier, '2' if they were less sure and '3' if they were very unsure. Subjects were tested individually, and each session lasted approximately 40 min. The significance level for all statistical tests was $p < .05$ unless otherwise stated.

Results

ROC analysis. Figure 3 presents the average ROCs for the semantic and perceptual encoding conditions plotted in probability space and *z* space. The ROCs for the semantic encoding condition were consistently higher than those for the perceptual encoding condition, showing that memory performance was better for the semantically processed items. The ROC for the perceptual processing condition exhibited a continuous inverted U shape in probability space and was fit well by a straight line in *z* space. This pattern has been reported in numerous other studies of recognition memory (e.g., Egan, 1958; Murdock, 1965; Ratcliff, Sheu, & Gronlund, 1992) and has been taken as support for the signal-detection-based models and as evidence against threshold models. However, the ROC for the semantically encoded items was only slightly curved in probability space and exhibited a slight U shape when plotted in *z* space. This suggests that the normality assumption underlying the signal-detection models was violated.

To facilitate comparison to previous studies the slopes and intercepts of the average z-ROCs were calculated. The slope and intercept for the perceptual encoding condition were .73 and 1.04, respectively. The respective values for the semantic encoding condition were .56 and 1.61. Note, however, that when the z-ROC is nonlinear the slope and intercept will depend on the specific response criteria that the subjects select.

Linear trend analysis supported the conclusions drawn from the visual inspection of the ROCs. For the ROCs in Fig. 3, linear equations were fit to the functions by minimizing the sum of squared errors (SSE). A quadratic term was then introduced to the equations to determine if there was a significant increase in the proportion of variance accounted for by introducing the nonlinear component. Because the ROCs were expected to be curved, and each point on the ROCs was free to vary in the *x* and *y* dimensions, the minimization of the SSE included variation in both hits and false alarms.

For the perceptually encoded items, the ROC in probability space was fit better by the nonlinear equation than by the linear equation ($F(1, 2) = 19.510$, $MSe = .0002$), showing that the ROC was curvilinear. Moreover, in *z* space, the ROC was fit well by a linear function, and there was no significant improvement associated with introducing the nonlinear component ($F(1, 2) = 1.756$, $Mse = .0004$). However, for the semantically encoded items, the ROC in probability space was fit reasonably well by a linear equation and was not fit significantly better by a nonlinear equation ($F(1, 2) = 5.515$, $Mse = .0001$), suggesting that the function was linear. Further, in *z* space the ROC was fit significantly better by the nonlinear than the linear equation ($F(1, 2) = 76.90$, $MSe = .0001$), showing that the z-ROC was curvilinear.

The results of the linear trend analysis prove problematic for both the threshold models and the signal-detection models. The ROC for the semantic encoding condition was relatively linear in probability space and U-shaped in *z* space, supporting the threshold models and contradicting the signal-detection models. However, the ROC for the perceptually encoded items was curved in probability space and was relatively linear in *z* space, supporting the signal-detection models and contradicting the threshold models.

Although the linearity analysis of the ROCs does not provide a direct assessment of the dual-process model, the observed functions were in general agreement with the model. The model predicts that the ROCs should exhibit a slight inverted U shape in probability space and a slight *U* shape in *z* space. This pattern was observed for the semantic processing ROC. For the perceptual processing condition the function in probability space was slightly curved, as expected, but the function in *z* space was fit well by a straight line. The failure to find a significant curve to the z-ROC might be taken as evidence against the dual-process model. However, such a conclusion would be premature, because the contribution of recollection was not expected to be large for the perceptually encoded items, and thus the deviation from linearity should

TABLE 1
Sum of Squared Error Terms Associated with the Best Fit for Each Model to the ROCs for the Semantic and Perceptual Encoding Conditions in Experiment 1

Number of parameters	Model	Encoding condition: Perceptual	Encoding condition: Semantic
1	High-threshold (R_o)	.01426	.00246
	2-Equal-threshold ($R_o = R_n$)	.01386	.07674
	Equal-variance signal-detection (d')	.00344	.00444
2	2-Threshold (R_o, R_n)	.00400	.00051
	Unequal-variance signal-detection (d', V_o)	.00011	.00029
	Dual process (d', R_o)	.00025	.00007

be quite small. The dual process model is assessed more directly in the next section, where it is fit to the observed data and compared to the alternative models.

Modeling the ROCs. The models were fit to the observed probability ROCs in Fig. 3 by reducing the sum of squared errors for each model. The idea was to find the simplest model that provided the best account for each ROC. The models were compared in the following way. First, the single-parameter models were examined. These included the high-threshold model, the 2-equal-threshold model, and the equal-variance signal-detection model. We will refer to these models as single-parameter models, because one parameter is sufficient to generate the ROC. However, note that additional parameters are required to fix individual points on the function. In the current experiment there were 5 points in each ROC, and thus 5 parameters were used to fix response criteria.

The best of the 1-parameter models was then compared to the best of the 2-parameter models. These included the 2-threshold model, the unequal-variance signal-detection model, and the dual-process model. If there was a significant decrease in the SSE between the 1- and 2-parameter models, then the 1-parameter models were rejected.

Table 1 shows the SSE terms for the semantic and perceptual encoding conditions for each model. For the perceptual condition, the dual process model and the unequal-variance signal-detection model provided the best accounts of the observed ROCs. The signal-detection model led to slightly better fit, but it did not differ significantly from the dual-process model. Of the single-parameter models, the equal-variance signal-detection model provided the best fit. However, of the 2-parameter models, the unequal-variance signal-detection model provided a significantly better fit than the single-parameter model ($F(1, 3) = 90.82$, $Mse = .0001$). To determine if there was a significant advantage for the unequal-variance signal-detection model over the dual-process model for the perceptual condition, the two models were fit to the ROCs for each subject. Because the SSE terms were not normally distributed, a Wilcoxon signed ranks test was used to compare the fits. The analysis showed that there was no reliable difference in how well the two models fit the observed ROC ($W = .588$, $p = .557$). Thus, for the perceptually encoded items, the dual-process and unequal-variance signal-detection models provided comparable accounts of observed ROC.

However, for the semantic processing ROC, the dual-process model provided a reliably better fit that any of the other models. For the 1-parameter models the high-threshold model provided the best fit. However, for the 2-parameter models, the dual-process model provided the best fit, and it reflected a significant improvement over the 1-parameter models ($F(1, 2) = 102.428$, $MSe = .00002$). Moreover, comparing the individual subject ROCs showed that the dual-process model provided a significantly better account of the ROC data than the unequal-variance signal-detection model ($W = 2.26$, $p = .012$).

A subsequent analysis was conducted to determine if ceiling effects had influenced the shape of the ROC in the semantic condition. Several subjects' ROCs in the semantic condition were close to ceiling (i.e., the hit rate approached 1.0), and it is possible that the advantage of the dual-process model over the signal-detection model was related to the inclusion of these subjects' ROCs in the analysis. However, when the data from five of the subjects whose ROCs came closest to ceiling was removed, the advantage of the dual-process model over the signal-detection model increased very slightly, suggesting that ceiling effects did not greatly influence ROC analysis.

The results of the modeling converge with the results of the linearity analysis, in showing that the recognition ROCs were in agreement with the dual-process model and that neither the signal-detection-based models nor the threshold-based models alone were sufficient to account for both ROCs. The curvilinearity of the ROC in the perceptual encoding condition argued against the threshold models, and the U shape of the z-ROC in the semantic encoding condition argued against the signal-detection models. Although the dual-process model and the unequal-variance signal-detection model provided comparable accounts of the perceptual encoding condition ROC, under semantic encoding conditions, where the two models made divergent predictions, the dual-process model provided the best fit.

EXPERIMENT 2

The results of Experiment 1 showed that the dual-process model provided an accurate account of the recognition ROC data. In Experiment 2 we tested the generalizability of those results by examining performance under slightly different study conditions. However, most importantly, Experiment 2 was designed to assess the phenomenological validity of the dual-process model, by contrasting the ROC data with subjective reports of recollection and familiarity. The idea was to ask subjects to report on the occurrence of recollection and familiarity using the remember/know procedure (Tulving, 1985) and to use these reports to predict the shape of the ROCs that would be obtained when the confidence data were plotted. If subjects are aware of recollection and familiarity and the dual process model is correct then we should be able to use their remember/know responses to predict the shape of their ROCs.

In the study phase, subjects heard words that were presented either once or twice. Although the study manipulation (once vs. twice) was not of direct interest, it allowed us to examine ROCs at two different levels of performance. The study words were presented by either a male or a female voice, and subjects were instructed to try to remember who said each word. The study conditions were similar to the semantic encoding conditions in the previous experiment in the sense that subjects were encoding information about each item that could be recollected at the time of test. Thus, the ROCs should reflect a sizable contribution of recollection as well as familiarity.

In the test phase, subjects were presented with a mixture of studied and new words and were required to make two memory judgments for each word. First, they were required to make a recognition judgment on a 6-point confidence scale. As in Experiment 1, the confidence data were used to plot ROCs and assess the memory models. Second, subjects were required to introspect about their memory judgments and to report whether they recognized an item on the basis of recollection or on the basis of familiarity. Using the remember/know procedure, we asked the subjects to respond R if they could recollect any aspect of the study event, K if they knew the item was studied but they could not recollect it, and N if they thought the item was new. Remembering was described to the subjects as the ability to consciously recollect some aspect of the study event, such as what happened or what was experienced during the presentation of the item at study. Subjects were instructed to respond that they knew the word was presented when they thought the word was studied but they could not recollect anything about the study event. That is, they should respond K if the item was familiar in the absence of recollection.

The remember/know data was used to estimate the contribution of recollection and familiarity. Given estimates of recollection and familiarity, the dual-process model is completely constrained and can be used to generate predicted ROCs. The predicted ROCs were then contrasted to the observed ROCs that were based on the subjects' confidence data. If subjects can reliably report on recollection and familiarity and the dual-process model is correct then we should be able to accurately predict the shape of their ROCs. To further contrast the remember/know and the ROC data, estimates for recollection and familiarity derived from the remember/know data were contrasted to estimates derived by fitting the dual-process model to the ROC data. The primary goal of the experiment was to examine the relationship between the ROC data and the subjective reports of recollection and familiarity and to determine if subjects were consciously aware of the processes underlying the dual-process model.

Method

Subjects and materials. Twenty-four students participated in the experiment for credit in an undergraduate psychology course. Two-hundred and forty medium to high-frequency words were selected from the Toronto word pool and randomly assigned to three lists of 80 words each.

Design and Procedure

During the study phase, two lists of 80 words each were presented auditorily using a cassette player at a rate of one word every 3 s. The first list was always read by a male voice and was presented twice in succession. The second list was read by a female voice and was presented only once. Subjects were instructed to try to remember which words were presented and which words were spoken by the male and the female voice.

Immediately after the study phase, subjects received a recognition memory test. Subjects were given a test booklet containing a randomized mixture of the 80 words that had been spoken by the male, 80 words spoken by the female, and 80 new words. There were two spaces beside each word for subjects to write their responses.

Subjects were told that they were to make two different memory judgments for each word. First, they were to rate on a 6-point confidence scale how sure they were that the word was presented in the study list. Subjects were instructed to write down a number from 1 (sure it was not studied) to 6 (sure it was studied) in the first space beside each word. Second, subjects were instructed to make remember/know judgments for each word. The remember/know instructions were based on those reported by Gardiner (1988). If they remembered the occurrence of the study word they were instructed to respond R, if they thought the word was studied but they could not recollect anything about the study event they were to respond K, and if they thought the word was new they were to respond N. Subjects were instructed that remembering could include recollecting who said the word, but it also included the recollection of any other aspect of the study event. Subjects were told to work through the booklet one word at a time, making a recognition and a remember/know judgment for each word before going on to the next. The experimental session took approximately 45 min. The significance level for all statistical tests was $p < .05$ unless otherwise stated.

Results

The remember/know judgments. In the remember/know task the probability of accepting an item as old (responding either R or K) was .75 for the items presented twice, .63 for the items presented once, and .33 for new items. A similar pattern was observed for items eliciting an R response. The probability of responding R was .46 for the items presented twice, .32 for the items presented once, and .05 for the new items.

Based on the remember responses we calculated the probability of recollection for the items studied once and twice. Because subjects falsely recollected some new items, the estimates for recollection incorporated performance on old and new items. In keeping with previous remember/know studies (e.g., Gardiner & Java, 1991) we simply subtracted the rate of false recollections from the rate of true recollections. This method of correcting for false recollection assumes a 2-equal-threshold model for recollection. Although we believe that a high-threshold model is more appropriate for recollection in this task, the specific model that is used to correct for false recollections does not make much of a difference, because the rate of false recollection in the current experiment was very low (.05).

Estimates of familiarity (d') were derived by solving the dual process equations presented earlier, based on the average hit rate, the false alarm rate, and the estimate of recollection, for each condition. For example, for the items presented once, the hit rate was .63, the false alarm rate was .33, and the estimate for recollection was .27 (.32–.05). Using these values we solved the dual-process equations to arrive at an estimate of familiarity ($d' = 0.42$). A simple way of solving the equations is to use d' look-up tables. For example, substituting the hit rate (.63) and the estimate of recollection (.27) into the dual-process equation for the hit rate leads to a value of .49 for ($F_o > c$). Using the false alarm rate (.33) and .49 as the hit rate, d' tables show the estimate of familiarity to be $d' = 0.42$.

Note that in contrast to most previous remember/know studies, our immediate interest was not in the proportion of K responses (the probability that an item was

TABLE 2
Estimates for Recollection and Familiarity Based on the Dual Process Analysis of the ROC Data and the Remember/Know Data for Experiment 2

		Estimation method	
Process	Study condition	Remember/know	ROC
Recollection	Once	.27	.28
	Twice	.41	.43
Familiarity	Once	0.42	0.40
	Twice	0.64	0.57

Note. Note that recollection is measured as a probability and familiarity is estimated as a d' value.

both familiar and not recollected (i.e., $F(1 - R)$); rather, we were interested in the probability that an item was familiar. Because subjects only respond K for the items that were not recollected, K will underestimate familiarity. Accordingly, familiarity can be calculated as the probability that subjects made a K response given that they had the opportunity to do so (i.e., $F = K/(1 - R)$). For a more detailed discussion of estimating familiarity using the remember/know procedure see Yonelinas and Jacoby (1995).

The estimates for recollection and familiarity based on the remember/know procedure are presented in Table 2. Based on these estimates, we used the dual-process model to generate predicted ROCs, and we contrasted those to the observed ROCs. The observed ROCs were plotted in the same way they were in Experiment 1 and are presented along with the predicted ROC functions in Fig. 4a. An examination of Fig. 4a shows that the observed ROC data was fit very well by the predicted functions. These results show that subjective reports of recollection and familiarity can be used along with the dual-process model to accurately predict the shape of the ROCs, thus providing support for the model.

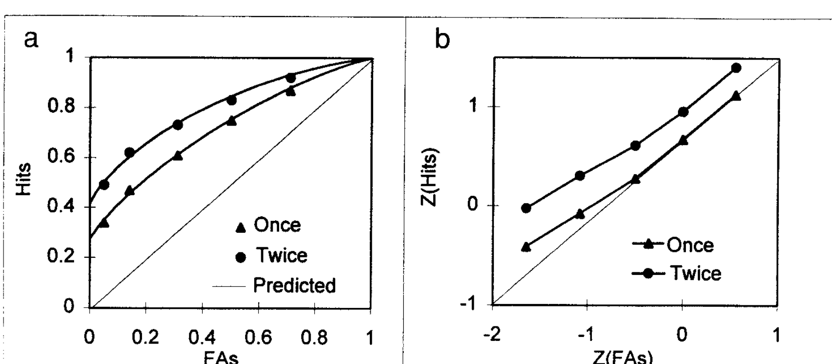

FIG. 4. (a) Predicted and observed recognition memory ROCs from Experiment 2 for items presented once and twice at study plotted on probability coordinates. (b) Observed ROCs plotted on z coordinates.

To further compare the remember/know and ROC data, estimates of recollection and familiarity were derived from the ROC data and compared to those derived from the remember/know data (see Table 2). The dual-process model was fit to the ROC data using the same curve-fitting procedure used in Experiment 1 (i.e., minimizing the SSE). The parameter values from the best fitting function were used as estimates of recollection and familiarity. An examination of Table 2 shows that the estimates derived from the ROC data were remarkably similar to those derived from the remember/know data, suggesting that the ROC results and the remember/know results reflect the same underlying processes.

The relation between ROCs and the subjective report of recollection was examined further by assessing the correlation between the estimates derived using the ROC analysis and those based on the remember responses. The recollection parameter derived from the ROC analysis was highly correlated (.84) with the measure derived from the remember/know procedure, $F(1, 46) = 106.70$, $MSe = .007$. These results suggest that at the level of individual subjects, the estimates for recollection derived using the ROC analysis parallel the estimates derived using the remember/know procedure.

The results of the remember/know and ROC analyses show that there is a strong relationship between ROCs and subjective reports of recollection and familiarity. The examination of remember/know responses allowed us to accurately predict the shapes of the ROCs. Moreover, the estimates of recollection and familiarity gained from the ROC analysis paralleled those based on subjects' introspective reports of those processes. These results make clear the relationship between ROCs and subjective reports of recollection and familiarity and show that the processes that underlie the dual-process model are available to conscious inspection.

The confidence and remember/know responses were examined further to assess one of the underlying assumptions of the dual-process model. The model assumes that recollection can be described as a high-threshold process and that relative to familiarity it leads to high confidence recognition responses. This means that estimates of recollection should remain constant as the recognition response criterion is varied. To assess these notions we examined the probability that studied items received a remember response as a function of recognition confidence. In agreement with the model, the estimates of recollection remained constant as response criterion was relaxed. For the items presented once at study the estimates of recollection, as the criterion was varied from strict to lax, were .30, .32, .32, .32, .32, and .32. The comparable estimates for items presented twice were .44, .46, .46, .46, .46, and .46. These results show that recollection did not vary with response criterion as might have been expected if it reflected a signal-detection processes.

ROC analysis. The ROC data was examined by assessing the linearity of the ROCs. As in Experiment 1 the analysis showed that the recognition data were in conflict with the threshold and signal-detection models. Figure 4 presents the average ROCs for the words presented once and twice, plotted in probability space and z space. The ROCs for the items presented once and twice exhibited a slight inverted U shape in probability space and a slight U shape in z space.

The results of a linear trend analysis were in agreement with these observations. For the items presented once at study, the ROC in probability space was fit better

TABLE 3
Sum of Squared Error Terms Associated with the Best Fit for Each Model to the ROCs for the Items Presented Once and Twice in Experiment 2

Number of parameters	Model	Encoding presentations: Once	Twice
1	High-threshold (R_o)	.00907	.01682
	2-Equal-threshold ($R_o = R_n$)	.00866	.02327
	Equal-variance signal-detection (d')	.00728	.00882
2	2-Threshold (R_o, R_n)	.00229	.00824
	Unequal-variance signal-detection (d', V_o)	.00045	.00040
	Dual process (d', R_o)	.00020	.00037

by the nonlinear equation than by the linear equation ($F(1, 2) = 65.51$, $MSe = .0001$), showing that the ROC was curvilinear. Moreover, in z space, the ROC was fit better by the nonlinear than the linear equation, ($F(1, 2) = 79.93$, $MSe = .0001$), showing that the z-ROC was also curvilinear. Similarly, for the items presented twice at study, the ROC in probability space was fit better by the nonlinear than the linear equation ($F(1, 2) = 121.69$, $MSe = .0001$), and the ROC in z space was fit slightly better by the nonlinear equation than the linear equation ($F(1, 2) = 15.19$, $MSe = .0004$, $p < .06$). The slope and intercept of the z-ROC for the items presented once were .70 and 0.69, respectively. The respective values for the items presented twice were .64 and 0.99.

Modeling the ROCs. The models were fit to the observed ROCs in the same way they were in Experiment 1. Table 3 presents the SSE terms for each model for the items presented once and twice at study. Examination of Table 3 shows that the dual-process model provided the best account of the ROCs for both study conditions. However, the unequal-variance signal-detection model provided fits that were only slightly poorer. For the items presented once, the equal-variance signal-detection model provided the best fit of the single-parameter models. However, of the 2-parameter models, the dual process model provided the best fit, and it was significantly better than that provided by the single-parameter models ($F(1, 3) = 106.20$, $MSe = .0001$). To determine if there was a significant advantage for the dual-process model over the unequal-variance signal-detection model, the models were fit to the ROCs of each subject. The analysis showed that there was no reliable difference in how well the two models fit the observed ROCs ($W = .543$, $p = .589$).

For the items presented twice, the equal-variance signal-detection model provided the best fit for the single-parameter models. However, for the 2-parameter models, the dual-process model provided the best fit, and it reflected a significant improvement over the 1-parameter models ($F(1, 2) = 68.51$, $MSe = .0002$). Moreover, the unequal-variance signal-detection model provided a fit that was slightly poorer than that of the dual-process model, but the difference was not statistically reliable ($W = .23$, $p = .819$).

The results of the ROC analyses were similar to those of Experiment 1. In agreement with the dual-process model, the ROCs exhibited an inverted U shape in probability space and a U shape in *z* space. The results of the linearity analysis are problematic for both the threshold and the signal-detection models. The inverted U shape of the ROCs in probability space lends support to the signal-detection models but contradicts the threshold models. However, the U shape of the ROCs in *z* space lends support to the threshold models but contradicts the signal-detection models. Although the unequal-variance signal-detection model provided a fit of the data that was close to that provided by the dual-process model, the other signal-detection and threshold models were rejected.

DISCUSSION

ROC analysis has been used to test theories of memory for more than 20 years. The current study shows that such an analysis is extremely useful in discriminating between alternative models. Subjective reports of recollection and familiarity (Tulving, 1985) have also been used quite extensively in the recent past to examine the processes underlying recognition memory (see Rajaram & Roediger, in press). However, until now the two literatures have remained quite separate. The current study shows that there is a direct relationship between these two areas of research by showing that the different processes that underlie the observed ROCs are available to subjective awareness.

The ROC data were in agreement with the predictions of the dual-process model in showing that the contribution of recollection to recognition memory performance led the z-ROCs to exhibit a U shape. For perceptually encoded items in Experiment 1, where the contribution of recollection was expected to be minimal, the ROC exhibited an inverted U shape in probability space and was relatively linear in *z* space. However, for the semantically encoded items, where the contribution of recollection was expected to be much larger, the ROC was relatively linear in probability space and was U-shaped in *z* space. Similarly in Experiment 2, the ROCs for items presented once and twice at study exhibited an inverted U shape in probability space and a U shape in *z* space.

Recognition memory ROCs are often fit reasonably well by linear functions when they are plotted in *z* space (e.g., Donaldson & Murdock, 1968; Murdock, 1965; Ratcliff, Sheu, & Gronlund, 1992), and thus the finding that the z-ROCs were U-shaped was somewhat surprising. However, similar results have been reported previously. For example, we have examined memory ROCs under conditions where performance was expected to rely primarily on recollection and found that the z-ROCs were consistently U-shaped. For example, tests of source memory, where subjects must discriminate between items from different sources, seem to require recollection of qualitative information about the study event. As expected if source discrimination relied on recollection, the resulting z-ROCs were found to be U-shaped (Yonelinas, submitted for publication). Furthermore, tests of associative recognition, where subjects must discriminate between intact and rearranged word pairs, also seems to require recollection, and the z-ROCs in this task were also found to be U-shaped (Yonelinas, in press).

The observed ROCs were problematic for both the threshold models and the signal-detection models. The curvilinear ROCs contradicted the threshold models by showing that the relationship between hits and false alarms was not linear. However, the curvilinear z-ROCs contradicted the signal-detection-based models by showing that the normality assumption that is central for those models was consistently violated. Although the current analysis was not exhaustive in its examination of alternative models—more complex threshold- or signal-detection-based models may provide better accounts of the data—the results are important in showing that the most common models fail to account for the observed ROCs.

In contrast to the threshold and signal-detection models, the dual-process model provided an accurate and parsimonious account of the observed ROC data. The model assumes that recollection and familiarity contribute to recognition memory performance, and that these processes can be described with two memory parameters (R and d', respectively). In every condition, the dual-process model provided a better account of the ROC data than the single-parameter threshold and signal-detection models. Of the other 2-parameter models the unequal-variance signal-detection model provided an account for the data that came the closest to that of the dual-process model, but in the semantic encoding condition, where the two models predicted significantly different ROCs, the dual-process model provided a more accurate account of the data.

Further support for the dual-process model was provided by the results of the remember/know analysis in Experiment 2. The subjective reports of recollection and familiarity were used to accurately predict the observed ROCs. Moreover, the estimates derived from the ROC data using the dual-process model paralleled those derived using the remember/know procedure. The convergence of the results from the ROC and the remember/know analysis suggests that the parameters underlying the dual-process model (R and d') are psychologically real, in the sense that they reflect processes that subjects are aware of and can consistently report on. These results are important because they show that the nature of the observed ROCs is directly related to the contribution of conscious recollection and familiarity.

The current ROC analysis showed that the dual-process and the unequal-variance signal-detection models can produce ROCs that are remarkably similar. In many cases it was not possible to discriminate between the two models. The reason for this is that when both recollection and familiarity contribute to performance, the U shape in the z-ROC that is predicted by the dual-process model can be extremely subtle. Although the U shape became apparent in all but the shallow processing conditions in Experiment 1, it was only when there was a large contribution of recollection (i.e., deep processing in Experiment 1) that the U shape was pronounced enough to discriminate between the dual-process and unequal-variance signal-detection models.

Given that the dual-process model predicts only a slight U shape to the z-ROC, it is not surprising that previous studies have often found that z-ROCs are fit reasonably well by straight lines. Moreover, averaging artifacts inherent in ROC methods may make it even more difficult to observe U-shaped z-ROCs. For example, Ratcliff, McKoon, and Tindell (1994) have shown that adding noise to recognition confidence ROC data tends to make the ROC exhibit an exaggerated inverted U shape. They showed that adding as little as 5% noise was sufficient to lead to a noticeable distortion of the ROC. Similarly, adding noise to a U-shaped z-ROC can make the z-ROC appear more linear. Given that large numbers of observations are required to derive ROCs, some proportion of the responses probably do reflect noise. However, given that several of the current z-ROCs were significantly U-shaped suggests that noise does not always completely eliminate the U shape of the z-ROCs.

The success of the dual-process model in accounting for the ROC and remember/know data suggests that its underlying assumptions are reasonable. However, there may be cases under which the model's assumptions do not hold, and thus it is useful to consider these assumptions more closely. One important assumption underlying the model is that recollection and familiarity are independent. The evidence in support of the independence assumption has been discussed at length elsewhere and we will not review it here. However, in brief, the arguments in favor of the independence assumption include: (1) Systematic process dissociations. For example, using the process dissociation procedure to estimate the contribution of conscious recollection and automatic influences of memory, it is only if the processes are assumed to be independent that manipulations associated with controlled processes are found to influence recollection and leave the automatic influences unaffected (for a review see Jacoby, Yonelinas, & Jennings, in press). (2) Systematic task dissociations. Dissociations between performance on direct and indirect tests of memory observed in numerous experimental and neuropsychological studies (for reviews, see Richardson-Klavehn & Bjork, 1988; Roediger & McDermott, 1993) suggest that these two types of task tend to tap independent memory processes. (3) Tests of the dual-process recognition model. Several of the dual-process model's qualitative predictions have been verified, and quantitative assessment of the model has shown that it provides an accurate account of the recognition data (Yonelinas, 1994; Yonelinas, submitted for publication; Yonelinas, in press; Yonelinas & Jacoby, 1996; Yonelinas & Jacoby, 1995). The current results join this later set of studies in providing support for the model and its assumption of independence.

As an alternative to the independence assumption one could assume that the two processes are mutually exclusive; however, exclusivity based models run into several problems. For example, Gardiner (1988) has assumed that the processes underlying recognition memory are mutually exclusive (also see Gardiner & Parkin, 1990) and that both recollection and familiarity are threshold processes. Accordingly, Gardiner and colleagues use the probability of remember and know responses as estimates of the underlying memory processes and use standard threshold methods of subtracting incorrect responses from correct responses to attain bias-free measures of these memory processes. The curvilinear ROCs observed in the current study suggests that the threshold assumption (at least with respect to familiarity) is incorrect. Moreover, ROC data prove problematic for exclusivity-based models. The problem is that estimates of familiarity derived using the exclusivity assumption lead to improbable conclusions and patterns of results that are extremely unstable. For example, Yonelinas and Jacoby (1995) examined the effects of changing the size of shapes between study and test (size congruency) on recognition memory. They used the R/K procedure, but had subjects rate the confidence of their K responses. The probability of a K response was then plotted as a function of confidence (i.e., an ROC). They found that when a strict response criterion was adopted, there were more K responses to studied than nonstudied shapes; however, as the response criterion was relaxed the

ROC associated with K responses dropped below the negative diagonal, paradoxically suggesting that studied shapes were less familiar than new shapes. Moreover, adopting a strict response criterion showed that size-congruent shapes led to more K responses than size-incongruent shapes, but as the response criterion was relaxed the effect reversed. In contrast, when the processes were assumed to be independent, studied shapes were found to be more familiar than new shapes, size-congruent shapes were found to be more familiar than size-incongruent shapes, and these patterns remained constant across all levels of response confidence.

As another alternative, one could assume that recollection is redundant with familiarity. That is, recollection does not contribute to overall recognition memory performance beyond what is already contributed by familiarity. It is often a redundancy model that is implicitly adopted in investigations of memory for source (also see Joordens & Merickle (1993), who proposed a redundancy model for a word-stem completion task). In a typical source memory experiment subjects study items from two difference sources. They are then given a recognition test in which they must first discriminate studied items from nonstudied items and then are asked to judge the source of the recognized items. A natural way to think of performance in this task is to assume that recognition judgments measure one type of memory (e.g., familiarity) and the source judgments measure another (e.g., recollection). Presumably, subjects must recognize an item before its source can be recollected; thus, items whose source can be recollected form a subset of those that can be recognized. Such a redundancy model holds that recollection is always accompanied by familiarity—an assumption opposite to the exclusivity assumption.

A redundancy model that is often used in studies of source memory is based on the multinomial model of Batchelder and Riefer (1990). The multinomial model, however, assumes the recognition and source memory reflect threshold processes. The current results show that the threshold assumption is inappropriate for recognition memory (see Kinchla, 1994, for a similar argument). Moreover, the redundancy model assumes that overall recognition performance reflects a single memory process that is not influenced by recollection. This is in conflict with the experimental and neuropsychological data suggesting that recognition reflects at least two separate processes (e.g., Aggleton & Shaw, 1996; Atkinson & Juola, 1974; Huppert & Piercy, 1976; Jacoby & Dallas, 1981; Mandler, 1980; Verfaellie & Treadwell, 1993). The model also conflicts with the current ROC data. For example, if recollection is redundant in recognition then contrary to the current findings, recollection should not influence the shape of the recognition ROCs.

Much more difficult than rejecting the redundancy and exclusivity models is rejecting more complex or ‘‘partially correlated’’ dual-process models. For example, it is possible that recollection and familiarity are positively correlated but not perfectly so. If the two processes can be shown to exhibit a consistent correlation, this could be modeled by introducing an additional correlation parameter into the independence model. However, in light of the evidence in favor of the independence assumption and the model’s success in accounting for the current data, complicating the model would seem to be premature.

A second important assumption for the dual-process model is that recollection is well described as a threshold process. We will first try to clarify what is meant by a threshold process and then consider conditions under which recollection may not behave like a classical threshold process. It has recently been argued that when one assumes that recollection is a threshold process that this entails that recollection is all-or-none in the sense that either all of the information about a study event is recollected or none of the information is recollected (Dodson & Johnson, 1996). However, this is not a requirement of threshold theories, and the dual-process model does not assume that subjects either recollect all of the study information or none of the study information. Threshold theory assumes that memory is all-or-none in the sense that the underlying memory distributions are discrete, rather than continuous, as in signal-detection theory (see Murdock, 1974). This does not rule out the possibility that memory is associated with different levels or strengths. In fact, the term ‘‘threshold’’ refers to the fact that an item will only be recognized if its memory exceeds a specific threshold; items falling above the threshold are recognized and items falling below the threshold are not. Similarly, if recollection is well described by a threshold model, this does not mean that subjects are only able to remember either everything or nothing about a study event. In fact, it is clear that subjects can recollect different aspects of study events under different conditions and further that under some conditions, recollection of different aspects of a study event are functionally independent (Yonelinas, 1994; Yonelinas & Jacoby, 1996).

The success of the dual-process model in accounting for the current ROC findings suggests that recollection was well described as a threshold process. Moreover, in agreement with the notion that recollection is associated with high confidence responses, the probability of remembering in Experiment 2 was found to remain constant as the response criterion was relaxed. However, future experiments may show that there are conditions under which recollection behaves in a more continuous manner. In such cases it may be necessary to extend the dual-process model by assuming that both familiarity and recollection reflect signal-detection processes (see Macmillan & Creelman for a discussion of multidimensional signal-detection models). This would necessarily complicate the model by requiring an additional free parameter to reflect the recollection response criterion.

Although future experiments will likely lead to modifications of the dual-process model, the current model does provide an accurate account of the existing ROC data and represents a useful model for understanding data from remember/know experiments. We take the current results as providing strong support for the dual-process model and for showing the importance of separating conscious recollection from familiarity when examining human memory.

ACKNOWLEDGMENTS

We thank William Banks and an anonymous reviewer for valuable comments on earlier versions of this article.

REFERENCES

Aggleton, J. P., & Shaw, C. (1996). Amnesia and recognition memory: A re-analysis of psychometric data. *Neuropsychologia,* **34,** 51–62.

Atkinson, R. C., & Juola, J. F. (1974). Search and decision processes in recognition memory. In D. H.

Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology, Vol. 1: Learning, memory & thinking.* San Francisco: Freeman.

Banks, W. P. (1970). Signal detection theory and human memory. *Psychological Bulletin,* **74,** 81–99.

Batchelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review,* **97,** 548–564.

Bayen, U. J., Murnane, K., & Erdfelder, E. (1996). Source discrimination, Item Detection, and Multinomial Models of source monitoring. *Journal of Experimental Psychology: Learning, Memory, & Cognition,* **22,** 197–215.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology,* **33A,** 497–505.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior,* **11,** 671–684.

Dodson, C. S., & Johnson, M. K. (1996). Some problems with the process-dissociation approach to memory. *Journal of Experimental Psychology: General,* **2,** 181–194.

Donaldson, W., & Murdock, B. B. (1968). Criterion change in continuous recognition memory. *Journal of Experimental Psychology,* **76,** 325–330.

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

Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory and Cognition,* **16,** 309–313.

Gardiner, J. M., & Java, R. I. (1991). Forgetting in recognition memory with and without recollective experience. *Memory and Cognition,* **19,** 617–623.

Gardiner, J. M., & Parkin, A. J. (1990). Attention and recollective experience in recognition memory. *Memory and Cognition,* **18,** 579–583.

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

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Hintzman, D. L. (1986). ''Schema abstraction'' in a multiple trace memory model. *Psychological Review,* **94,** 341–358.

Huppert, F., & Piercy, M. (1976). Recognition memory in amnesic patients: Effects of temporal context and familiarity of material. *Cortex,* **12,** 3–20.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General,* **3,** 306–340.

Jacoby, L. L., Woloshyn, V., & Kelley, C. M. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by divided attention. *Journal of Experimental Psychology: General,* **118,** 115–125.

Jacoby, L. L., Yonelinas, A. P., & Jennings, J. M. (in press). The relationship between conscious and unconscious (automatic) influences: A declaration of Independence. In J. Cohen & J. W. Schooler (Eds.). *Scientific approaches to the question of consciousness.* Hillsdale, NJ: Earlbaum.

Johnson, M. K., Kounios, J., & Reeder, J. A. (1994). Time-course studies of reality monitoring and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **20,** 1409–1419.

Kinchla, R. A. (1994). Comments on Batchelder and Riefer's multinomial model for source monitoring. *Psychological Review,* **101,** 166–171.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown Univ. Press.

Kunst-Wilson, W. R., & Zajonc, R. B. (1980). Affective discrimination of stimuli that cannot be recognized. *Science,* **207,** 1019–1024.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A users guide.* Cambridge Univ. Press, Cambridge.

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

Murdock, B. B. (1965). Signal-detection theory and short-term memory. *Journal of Experimental Psychology,* **70,** 443–447.

Murdock, B. B. (1974). *Human memory: Theory and data.* Hillsdale, NJ: Earlbaum.

Murdock, B. B. (1982). A theory of storage and retrieval of item and associative information. *Psychological Review,* **89,** 609–66.

Rajaram, S., & Roediger, H. L. (in press). Remembering and knowing as states of consciousness during recollection. In J. D. Cohen & W. J. Schooler (Eds.), *Scientific approaches to the study of consciousness.* New York: Erlbaum.

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, & Cognition,* **20,** 763–785.

Ratcliff, R., Sheu, C. F., & Gronlund, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review,* **3,** 518–535.

Richardson-Klavehn, A., & Bjork, R. A. (1988). Measures of memory. *Annual Review of Psychology,* **39,** 475–543.

Roediger, H. L., & McDermott, K. B. (1993). Implicit memory in normal human subjects. In H. Spinnler, & F. Boller (Eds.), *Handbook of neuropsychology* (Vol. 8). Amsterdam: Elsevier.

Swets, J. A. (1986). Indices of discrimination or diagnostic accuracy: Their ROCs and implied models. *Psychological Bulletin,* **99,** 100–117.

Swets, J. A., Tanner, W. P., & Birdsall, T. G. (1961). Decision processes in perception. *Psychological Review,* **68,** 301–340.

Toth, J. P. (1996). Conceptual fluency in recognition memory: Levels of processing effects on familiarity. *Canadian Journal of Experimental Psychology,* **50,** 123–138.

Tulving, E. (1985). Memory and Consciousness. *Canadian Psychologist,* **5,** 1–13.

Verfaellie, M., & Treadwell, J. R. (1993). The status of recognition memory in amnesia. *Neuropsychology,* **1,** 5–13.

Warrington, E. K., & Weiskrantz, L. (1974). The effect of prior learning on subsequent retention in amnesic patients. *Neuropsychologia,* **12,** 419–428.

Wickelgren, W. A., & Norman, D. A. (1966). Strength models and serial position in short-term recognition memory. *Journal of Mathematical Psychology,* **3,** 316–347.

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

Yonelinas, A. P., & Jacoby, L. L. (1995). The relation between remembering and knowing as a basis for recognition: Effects of size congruency. *Journal of Memory and Language,* 34, 622–643.

Yonelinas, A. P., & Jacoby, L. L. (1996). Noncritical Recollection: Familiarity as Automatic, Irrelevant Recollection. *Consciousness and Cognition,* **5,** 131–141.

Yonelinas, A. P. Recognition memory ROCs for item and associative information: Evidence for a dual process signal detection model. *Memory and Cognition,* in press.

Yonelinas, A. P. The contribution of recollection and familiarity to recognition and source memory, submitted for publication.

Received July 22, 1996