Psychological Review
2004, Vol. 111, No. 2, 524–542

0033-295X/04/$12.00 DOI: 10.1037/0033-295X.111.2.524

Remember–Know: A Matter of Confidence

John C. Dunn
University of Western Australia

This article critically examines the view that the signal detection theory (SDT) interpretation of the remember–know (RK) paradigm has been ruled out by the evidence. The author evaluates 5 empirical arguments against a database of 72 studies reporting RK data under 400 different conditions. These arguments concern (a) the functional independence of remember and know rates, (b) the invariance of estimates of sensitivity, (c) the relationship between remember rates and overall hit and false alarm rates, (d) the relationship between RK responses and confidence judgments, and (e) dissociations between remember and overall hit rates. Each of these arguments is shown to be flawed, and despite being open to refutation, the SDT interpretation is consistent with existing data from both the RK and remember–know–guess paradigms and offers a basis for further theoretical development.

The remember–know (RK) paradigm was introduced by Tulving (1985) in order to measure the different states of awareness thought to underlie memory retrieval. Participants in a memory experiment were asked to indicate the basis on which they judged that an item had been previously studied. They were asked if they were able to "remember" its prior occurrence or whether they simply "knew" on some other basis that it was old. Tulving (1985) interpreted remember (or R) responses as indicating a state of *autonoetic awareness* associated with retrieval from episodic memory and interpreted know (or K) responses as indicating a state of *noetic awareness* associated with retrieval from semantic memory.

The RK paradigm was further refined by Gardiner (1988), who developed operational definitions of both remembering and knowing. *Remembering* was defined as the ability to become consciously aware of some aspect or aspects of what occurred or was experienced at the time the test item was first presented. *Knowing* was defined as recognition that the test item had been presented earlier but without the ability to recollect consciously anything about its actual occurrence or what had happened or was experienced at that time. In addition, Gardiner (1988) reported a dissociation between R and K responses. He found that levels of processing manipulations affected the proportion of R responses but had no effect on the proportion of K responses.

Since these two pioneering studies, there has been growing interest in the RK paradigm (Gardiner & Richardson-Klavehn, 2000). Much of this research has investigated the extent to which R and K responses can be dissociated by different experimental variables. The results are clear—R and K responses are functionally dissociable such that "all possible relations between these two states of awareness have been observed" (Gardiner, Ramponi, & Richardson-Klavehn, 1998, p. 2). It is thus possible to identify variables that affect R responses but not K responses, variables that affect K responses but not R responses, variables that affect R and K responses in the same direction, and variables that affect these responses in opposite directions (Gardiner, 2001).

Studies that use the RK paradigm frequently do so under the assumption that R and K responses reflect different forms of memory retrieval. On this view, and consistent with Tulving's original proposal, R and K responses are interpreted as reflecting the operation of two qualitatively different memory components, systems, or processes. This view is here called the *dual-process interpretation* of the RK paradigm. It subsumes a number of competing models concerning the nature of the underlying processes, the kinds of variables that affect them, and how they in turn affect the likelihood of R and K responses. In a recent review, Gardiner (2001) identified three such models. The first of these is Tulving's proposal that R responses reflect the subjective experience of retrieval from episodic memory whereas K responses reflect the subjective experience of retrieval from semantic memory. The second is a proposal by Rajaram (1996) that R responses reflect the distinctiveness of processing at study whereas K responses reflect the fluency of processing at test. The third is a proposal by Jacoby, Yonelinas, and Jennings (1997) that R and K responses can be identified with the processes of *recollection* and *familiarity*, respectively, that are thought to underlie recognition memory (for a review of this framework, see Yonelinas, 2002).

In contrast to the dual-process interpretation of the RK paradigm, several researchers have proposed that R and K responses reflect different levels of confidence concerning the products of memory retrieval (e.g., Donaldson, 1996; Hirshman, 1998; Inoue & Bellezza, 1998). On this view, the instructions to respond "remember" or "know" are interpreted by participants as requiring them to adopt more and less stringent criteria, respectively, for recognition or recall. During recognition, if the evidentiary or so-called trace strength of a test item exceeds the more stringent criterion, an R response is made; if it is between the two criteria, a K response is made; and if it is less than the less stringent criterion, a "new" response is made. Because this interpretation of the RK paradigm is couched in the language and terminology of

I thank Kim Kirsner for his helpful comments on drafts of the article. The remember–know database referred to in this article is available on the Internet at www.general.uwa.edu.au/u/kraepeln/people/jdunn/rkdatabase.xls.

Correspondence concerning this article should be addressed to John C. Dunn, School of Psychiatry and Clinical Neurosciences, University of Western Australia, 35 Stirling Highway, Nedlands, WA 6009, Australia. E-mail: jdunn@cyllene.uwa.edu.au

the theory of signal detection, it is called the *signal detection theory* or *SDT interpretation* of the RK paradigm.

It is important to distinguish the dual-process and SDT interpretations of the RK paradigm for three main reasons. First, they offer radically different frameworks for interpreting the results of the RK paradigm. Because competing theoretical accounts of R and K responses are predicated on one or other interpretation, it is first necessary to establish which, if either, is valid. Second, Hirshman and Master (1997) have argued that formal testing of the SDT interpretation would allow researchers to determine which findings from the RK paradigm provide strong evidence for the existence of two separate memory systems or processes. On this view, the SDT interpretation serves as a kind of null hypothesis against which the validity of the dual-process interpretation can be assessed. Thus, Hirshman, Lanning, Master, and Henzler (2002) proposed that “one critical purpose of the [SDT] model is through appropriate falsification to allow stronger inferences about the nature of consciousness and recognition memory processes” (p. 153). Third, SDT offers a framework for understanding human memory that has been successfully applied over a long period that continues to the present (e.g., Banks, 1970; DeCarlo, 2002; Murdock, 1974). Given the utility of this approach, its rejection in favor of an alternative should be based on the strongest evidence.

As Hirshman and his colleagues have argued, the SDT interpretation may be viewed as a kind of null hypothesis in relation to other more complex accounts. In general, because the dual-process interpretation posits the existence of two qualitatively different forms of memory, whereas the SDT interpretation posits only one, dual-process models tend to be more complex than models based on the SDT interpretation. As a consequence, it should be possible to find data that, while consistent with dual-process interpretation, are inconsistent with the SDT interpretation. To this end, a number of studies have sought to identify such data and to show thereby that the SDT interpretation is incorrect. In a recent review of several of these studies, Gardiner (2000, p. 1356) concluded that “this trace strength model [the SDT interpretation] does not fit the data, except under exceptional circumstances.” The aim of the present article is to show that this view is mistaken. In point of fact, rather than there being compelling evidence against the SDT interpretation, the existing data are completely consistent with this account.

The present article examines five empirical arguments that have been used to refute the SDT interpretation. Some of these arguments have been summarized by Gardiner (2001) and Yonelinas (2002) and have also been recently discussed by Conway, Dewhurst, Pearson, and Sapute (2001); Dobbins, Khoe, Yonelinas, and Kroll (2000); Donaldson (1996); Gardiner (2000); Gardiner and Richardson-Klavehn (2000); Gardiner, Ramponi, and Richardson-Klavehn (1998); and Gardiner and Conway (1999), among others. Each of the arguments has the same basic logical form consisting of two premises and a conclusion. The first premise is a claim that the SDT interpretation is inconsistent with a particular empirical outcome. The second premise is a demonstration that the outcome in question can be observed. The conclusion is that on the basis of these premises being true, the SDT interpretation can be rejected. In each of the five arguments to be examined, the first premise will be shown to be false. As a consequence, because the SDT interpretation is not inconsistent with the observed outcome, it cannot be used to reject the account.

The remainder of the present article consists of the following sections. First, the signal detection interpretation of the RK paradigm is formally presented in terms of a set of defining equations. By assuming a particular form for these equations, one instantiation of the SDT interpretation, the equal-variance normal distribution model, is also defined. Second, an observed database consisting of the results of 72 studies and 400 different experimental conditions is described. The purpose of this database is to determine whether an empirical outcome, proposed as being ruled out by the SDT interpretation, does or does not exist. A model-specific database is also defined, which consists of the set of predictions of the observed database derived from the equal-variance normal distribution model. The purpose of this database is to determine whether a given empirical outcome is or is not ruled out by the SDT interpretation. The next five sections evaluate each of the empirical arguments in turn. Following this, I examine the question of whether the SDT interpretation is able to provide a psychologically plausible account of the data. Although much of the relevant work is lacking, it can be demonstrated that the underlying signal detection parameters are affected in meaningful ways by different independent variables. In the final section, I conclude that the SDT interpretation offers a viable, simple, and meaningful alternative account of the RK paradigm.

The SDT Interpretation

The SDT interpretation of the RK paradigm consists of three core assumptions (Hirshman & Master, 1997). The first assumption is that items in a memory task can be ordered along a single continuum of familiarity or strength of evidence. The second assumption is that the strength of evidence of old or studied items is on average greater than the strength of evidence of new or nonstudied items. That is, the strength-of-evidence distribution of old items can be considered as being displaced by a certain amount relative to the strength-of-evidence distribution of new items. The third assumption is that the proportions of R and K responses are determined by the placement of two criteria on the strength-of-evidence axis. The less stringent criterion is used to distinguish old (R and K) from new responses. The more stringent criterion is used to distinguish R from K responses.

The SDT interpretation can be formally expressed in terms of the following set of equations. Let R_O be the probability of an R response to an old item and let R_N be the probability of an R response to a new item. Similarly, let K_O be the probability of a K response to an old item and let K_N be the probability of a K response to a new item. Then,

$$
\begin{aligned}
R_O &= 1 - F_O(r - d) \\
R_N &= 1 - F_N(r) \\
R_O + K_O &= 1 - F_O(k - d) \\
R_N + K_N &= 1 - F_N(k), \qquad (1)
\end{aligned}
$$

where, F_O is the cumulative distribution function[1] for the set of old items, F_N is the cumulative distribution function for the set of new items, d is a parameter representing the displacement of the distribution of old items relative to the distribution of new items, r is the criterion separating R from K responses, and k is the criterion separating old from new responses. The quantities $R_O + K_O$ and $R_N + K_N$ are also known as the old–new or overall hit and false alarm rates, respectively. In the present formulation, the cumulative distribution functions of old and new items may be different, which allows for the possibility, for example, that the shape or variances of the two distributions may not be the same.

Because, in general, the functions F_O and F_N are unknown, it is not possible to derive quantitative predictions from Equation 1. For this reason, two additional distributional assumptions are commonly made. These are (a) $F_O = F_N = F$ and (b) $F = \Phi$, the normal cumulative distribution function. The SDT interpretation, together with these additional assumptions, defines the following equal-variance normal distribution model:

$$R_O = 1 - \Phi(r - d')$$

$$R_N = 1 - \Phi(r)$$

$$R_O + K_O = 1 - \Phi(k - d')$$

$$R_N + K_N = 1 - \Phi(k). \tag{2}$$

In this case, the parameter d in Equation 1 corresponds to the standard measure of sensitivity, d'.

It is important to distinguish the more general SDT interpretation of the RK paradigm given by Equation 1 from the particular equal-variance normal distribution model given by Equation 2. Although the latter is useful in generating quantitative predictions, it does not exhaust the possible choices of the strength-of-evidence distributions of old and new items. Therefore, if this model is shown to be inconsistent with the data, it may mean either that the underlying SDT interpretation is incorrect or that this set of distributional assumptions is simply wrong. The primary value of the equal-variance normal distribution model is that it may be used as part of an existence proof. If it is claimed that a pattern of data is inconsistent with the SDT interpretation, then this pattern must also be inconsistent with the equal-variance normal distribution model. Therefore, if it can be shown that the pattern of data is consistent with this model, then the claim that it is inconsistent with the SDT interpretation is incorrect.

The Observed and Model-Specific Databases

In the following sections, predictions derived from the SDT interpretation are evaluated against an observed database consisting of the results of 72 studies that report data from the RK paradigm. These studies are listed in the Appendix. A study was included in the database if it met the following criteria: (a) It could be identified in a computer search of the relevant literature, (b) it reported the results of at least one experiment using the RK paradigm, (c) the data included rates of R and K responses for both old and new items, and (d) only R and K responses were allowed for old items (i.e., this database excludes studies that used an additional guessing category). The majority of included studies tested memory using a yes/no recognition task, although some studies used cued recall or forced-choice recognition tasks. In addition, some studies used the RK paradigm to study false or illusory memory. Of these, only the results for items actually presented have been included. Furthermore, because different studies tend to use different instructions to participants concerning R and K responses, a lenient standard was used.

Although every attempt was made to detect as many publications as possible, I make no claim that the observed database is exhaustive. However, it is likely to be reasonably comprehensive and to include a representative cross-section of RK research conducted since 1985. The final set of 72 studies includes the 17 studies originally assembled by Donaldson (1996)[2] and contains the results of 124 separate experiments involving a total of 400 different experimental conditions reflecting manipulation of a wide range of independent variables, including the following: variation in levels of processing at study; retrieval demands, such as focused and divided attention; retention interval; stimulus characteristics, such as auditory or visual presentation or pictures and words; the age of participants; the presence or absence of a memory disorder; and the administration of drugs and alcohol.

In addition to the observed database, a second model-specific database was also constructed. This database consists of the predicted proportions of R and K responses for old and new items obtained by fitting the equal-variance normal distribution model separately to each experimental condition in the observed database. Values for the parameters d', k, and r for each condition were derived by minimizing the sum of squared differences between the observed and predicted data. The model-specific database provides a means of verifying any claim that a particular pattern of data is inconsistent with the SDT interpretation. Because this database is, by construction, completely consistent with the SDT interpretation, any pattern of data that is proposed to be inconsistent with this interpretation cannot occur in the database. Therefore, if such a pattern can be found, it follows that it must be consistent with the SDT interpretation.

Argument 1: Functional Independence

The fact that R and K responses are functionally independent has been taken as lending support to the dual-process view. If R and K responses are based on different memory systems or processes, then it should be possible to find variables that selectively affect each such system or process. However, despite an impression to the contrary, functional independence is also consistent with the SDT interpretation.

Table 1 presents the results of four studies, each of which reported one of four qualitatively different effects of an experimental variable on the proportions of R and K responses. The first study, by Schacter, Verfaellie, and Anes (1997), examined the effect of presenting conceptually or perceptually related lists on false recognition in amnesics and controls. The results presented in Table 1 correspond to recognition of previously presented (true) targets in the conceptual condition of Experiment 1. These results

[1] Let X be a continuous real-valued random variable. Its cumulative distribution function is $f(x) = P(X \le x)$, where $P(e)$ is the probability of e.

[2] I am grateful to Wayne Donaldson for these data and an additional set of publications not included in Donaldson (1996).

Table 1
Fit of the Equal-Variance Normal Distribution Model to Data From Studies Showing Functional Independence of R and K Responses

Study and condition	Observed Old *R*	Observed Old *K*	Observed New *R*	Observed New *K*	Best-fitting model parameters d'	k	r	Predicted Old *R*	Predicted Old *K*	Predicted New *R*	Predicted New *K*
Schacter et al. (1997, Exp. 1)											
Amnesic	0.15	0.29	0.06	0.17	0.60	0.75	1.62	0.15	0.29	0.05	0.17
Control	0.45	0.28	0.03	0.08	1.85	1.25	1.98	0.45	0.28	0.02	0.08
Gregg & Gardiner (1994, Exp. 2)											
Auditory	0.10	0.27	0.05	0.09	0.86	1.19	2.12	0.11	0.27	0.02	0.10
Visual	0.11	0.52	0.03	0.18	1.21	0.88	2.43	0.11	0.52	0.01	0.18
Gardiner & Java (1990, Exp. 2)											
Word	0.28	0.16	0.04	0.11	0.95	1.07	1.55	0.28	0.18	0.06	0.08
Nonword	0.19	0.30	0.03	0.12	1.01	1.04	1.89	0.19	0.30	0.03	0.12
Gardiner et al. (1996, Exp. 1 & 2)											
1 trial	0.28	0.31	0.11	0.14	0.87	0.70	1.45	0.28	0.29	0.07	0.17
4 trials	0.38	0.45	0.02	0.09	2.25	1.30	2.56	0.38	0.45	0.01	0.09

Note. R = mean proportion of remember responses; K = mean proportion of know responses; d' = sensitivity; k = old–new decision criterion; r = remember–know decision criterion; Exp. = experiment.

show that in this condition, the presence of a memory deficit affected the proportion of R responses but had little or no effect on the proportion of K responses. The second study, by Gregg and Gardiner (1994), compared visual and auditory presentation of test items in an experiment in which old words had been presented earlier under conditions designed to minimize deep or elaborative processing. These words were presented visually for 300 ms with a 200-ms interstimulus interval, and participants were instructed to identify if any blurred letters were present. Changing the presentation modality from visual to auditory in the test phase decreased the proportion of K responses but had little or no effect on the proportion of R responses. The third study, by Gardiner and Java (1990), examined recognition memory for common words and pronounceable nonwords matched on length and number of syllables. This variable was found to have had opposite effects on the proportions of R and K responses— words led to more R responses whereas nonwords led to more K responses. Finally, the study by Gardiner, Kaminska, Dixon, and Java (1996) investigated memory for Polish folk tunes (Experiment 1) and classical music themes (Experiment 2). Participants heard each stimulus either once, twice, or four times. Because similar results were found in both experiments, the data have been averaged across experiment for greater clarity. This averaging reveals that an increase in the number of stimulus presentations led to an increase in the proportions of both R and K responses.

If the different patterns of association or dissociation between R and K responses are inconsistent with the SDT interpretation, then the equal-variance signal detection model should not be able to fit the data. The central three columns of Table 1 list the best-fitting parameter values of the equal-variance signal detection model for each of the four studies. The corresponding predicted values are given in the rightmost four columns. It is apparent that the predicted proportions of R and K responses are very similar to the observed proportions and that the relevant pattern of association or dissociation has been captured by the model. Thus, none of these patterns of data is inconsistent with the SDT interpretation.

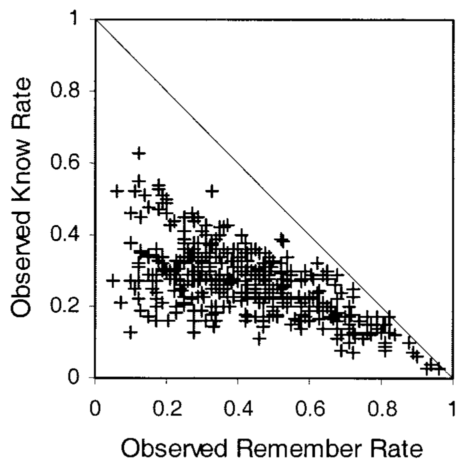
The four studies listed in Table 1 were selected as particularly clear examples of the different kinds of relationship between R and K responses that demonstrate functional independence. However, it might be argued that the SDT interpretation may fail to account for other kinds of relationship. That this is unlikely is shown by Figure 1. Figure 1A is a scatter plot of the mean proportion of K responses against the mean proportion of R responses for each entry in the observed database. Each point in this figure corresponds to one of the 400 conditions of the database. The functional independence of the two kinds of response is indicated by the fact that these variables are not perfectly correlated. It is possible to find pairs of conditions[3] that differ only on K responses, others that differ only on R responses, and others that differ on both R and K responses either in the same or opposite directions. Yet these patterns are also consistent with the SDT interpretation. Figure 1B presents the predictions of the best-fitting equal-variance signal detection model drawn from the model-specific database for each of the 400 conditions. Each point in this figure corresponds to the proportion of R and K responses for particular values of d', k, and r. It is apparent that it is also possible to find pairs of conditions that demonstrate the same relationships between R and K responses. These results demonstrate that the fact of functional independence of R and K responses is equally consistent with both the dual-process and SDT interpretations.

Argument 2: Estimates of Sensitivity

A second argument against the SDT interpretation is based on a proposition first advanced by Donaldson (1996), ironically in defense of this interpretation. The proposition is that, according to Equation 1, an estimate of sensitivity based on R responses should be equal to an estimate of sensitivity based on both R and K responses. This follows from the fact that the same parameter, d, appears in the equation for R_O as in the equation for $R_O + K_O$. Donaldson (1996) reviewed the results of 80 experimental condi-

[3] For present purposes, it is not necessary that pairs of conditions be drawn from the same experiment or reflect manipulation of an identifiable factor.

A



B

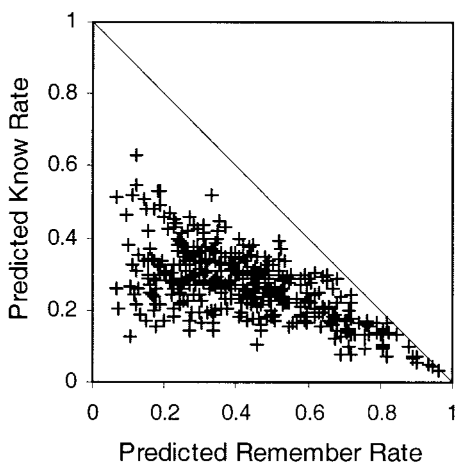


Figure 1. Relationship between the proportions of remember and know responses across 400 experimental conditions. A: Observed database. B: Model-specific database.

tions based on 17 studies and concluded that the two estimates were approximately equal. However, in a later analysis of an expanded version of this database consisting of 182 different conditions, Gardiner and Gregg (1997) found that "of the experimental conditions that show a difference . . . the [sensitivity] estimates derived from overall hit and false alarm rates are greater than those derived from remember hit and false alarm rates in 127 out of 162 cases" (p. 475). That is, although the two estimates of sensitivity are *approximately* equal, it is clear that there is a consistent trend for estimates based on R and K responses to be greater than those based on R responses alone.

In the analyses of both Donaldson (1996) and Gardiner and Gregg (1997), sensitivity was measured using the statistic A', although an analysis based on d' was also undertaken by Donaldson (1996). Given that the forms of the distributions of old and new items are unknown, considerable care is required when attempting to estimate model parameters. Drawing on the observation that A' is less affected by differences in the variances of the two distributions (Donaldson, 1993), it seemed appropriate to attempt to estimate sensitivity in the RK paradigm using both statistics, A' and d'. In addition, A' has a reputation for being a "nonparametric" index of sensitivity (Grier, 1971), which also suggests that it may be largely unaffected by the forms of the underlying distributions. In any event, A' has dominated analyses of estimates of sensitivity in the RK paradigm. It is defined as

$$A' = 0.5 + \frac{(h - f)(1 + h - f)}{4h(1 - f)}, \tag{3}$$

where h is the relevant hit rate and f is the corresponding false alarm rate.

The meta-analysis conducted by Gardiner and Gregg (1997) revealed that the estimate of A' based on R and K responses, A'_{RK}, was consistently greater than the estimate of A' based only on R responses, A'_R. This result is replicated in the current database. Across the 400 experimental conditions of the observed database, the mean value of A'_{RK} is 0.827 and the mean value of A'_R is 0.803. Although this difference is small, A'_{RK} exceeded A'_R in 307 or 76.8% of the conditions. This value accords well with the value of 78.4% found by Gardiner and Gregg (1997) and the value of 75% found by Donaldson (1996). Figure 2A shows the scatter plot of A'_{RK} against A'_R, revealing, as these statistics indicate, that the bulk of the observed data lie above the line of equality.

Although Gardiner and Gregg (1997) concluded that the observed discrepancy in the estimates of A' refuted the SDT interpretation, it is, in fact, completely consistent with this account. The reason for this is that A' is not invariant under different distributional assumptions (Macmillan & Creelman, 1996). In fact, it is a hybrid measure consisting of a combination of two measures of sensitivity, one based on the rectangular distribution, the other on the logistic distribution. The effect of this is that if the proportion of hits and false alarms are drawn from another distribution, A' will depend on both sensitivity *and* the decision criterion. This point is illustrated in Figure 3, which plots A' as a function of the decision criterion (c) for different values of sensitivity (d') under the assumptions that the distributions of old and new items are normal and have equal variance. That is, in terms of Equation 3, $h = 1 - \Phi(c - d')$ and $f = 1 - \Phi(c)$. If A' depends only on sensitivity then, for a given value of d', the function relating A' to c should be flat. However, none of the functions shown in Figure 3 have this property. Instead, each function is curvilinear, the exact shape of which depends on the choice of d'. In all cases, however, for both very high and very low values of c, A' converges on the value of 0.75. At intermediate values, it attains a local minimum or maximum at $c = d'/2$ (corresponding to the unbiased observer).

The fact that A' is affected by both sensitivity and the decision criterion under the equal-variance normal distribution model affects the interpretation of comparisons of A' across different values of c. This point is also illustrated in Figure 3. From the model-specific database, it is possible to calculate the mean values of the three parameters, k, r, and d', across all of the 400 data points. These values are 0.96, 1.75, and 1.53, respectively. The A' function corresponding to $d' = 1.53$ is shown in boldface in Figure 3, and the positions of the corresponding decision criteria, $c = 0.96$ and $c = 1.75$, are shown by the two vertical lines. One can see that the expected value of A' for $c = 0.96$, corresponding to the old–new criterion, is greater than the expected value of A' for $c = 1.75$, corresponding to the RK criterion. That is, under the equal-variance normal distribution model and for values of d' in the vicinity of 1.53, A'_{RK} should be greater than A'_R.

The model-specific database consists of the best-fitting estimates of the data under the equal-variance normal distribution model. It follows from the foregoing analysis that A'_{RK} should be systematically greater than A'_R across the 400 conditions in this database. As expected, A'_{RK} exceeded A'_R on 351 or 87.8% of the conditions. For these data, the mean value of A'_{RK} was 0.829 and the mean value of A'_R was 0.805. These values are in close accord with the corresponding values in the observed database, 0.827 and 0.803, respectively. Thus, rather than predicting that A' should be the same whether based on both R and K responses or only on R responses, the equal-variance normal distribution model predicts that the former should consistently exceed the latter, in close accordance with the observed data.

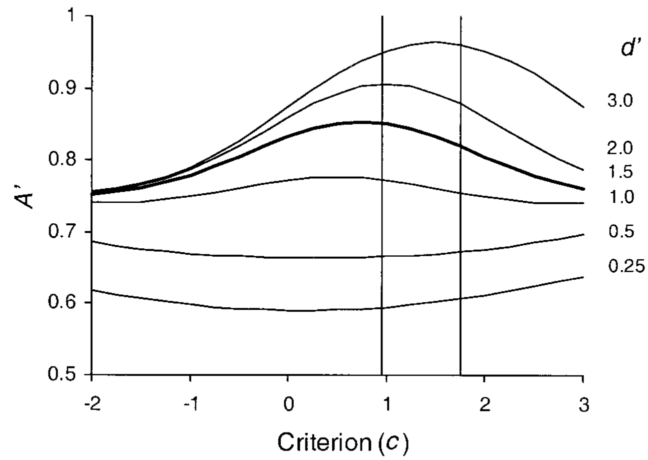


Figure 3. Relationship between A' and criterion (c) for different values of sensitivity (d') under the equal-variance normal distribution signal detection model. Vertical lines correspond to values of $c = 0.96$ and $c = 1.75$.

A

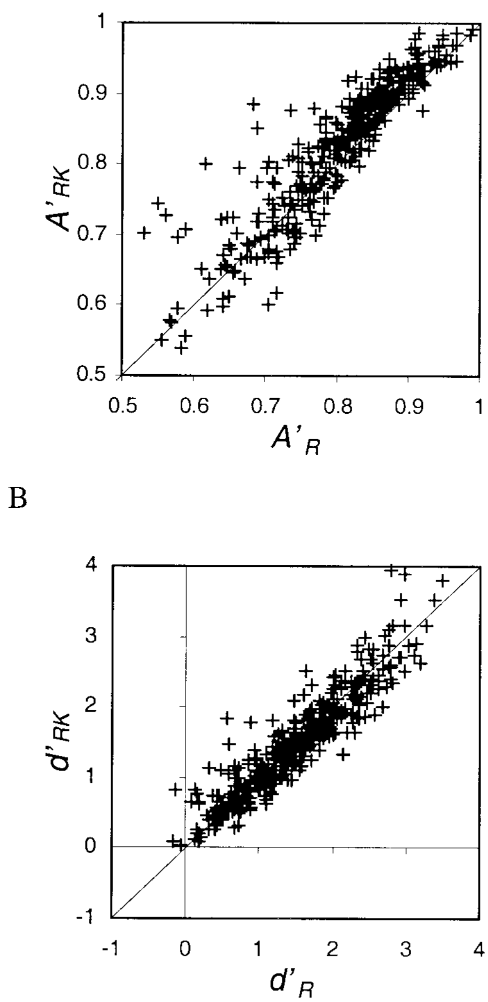


Figure 2. The relationship between sensitivity based on old–new hit rate, remember–know (RK), and sensitivity based on remember rate (R) across 400 conditions in the observed data set. A: Sensitivity defined by A'. B: Sensitivity defined by d'.

As well as comparing A' values derived from R responses with those derived from R and K responses, Donaldson (1996) also conducted a similar analysis using d'. The mean value of d' based on the sum of R and K responses, d'_{RK}, was found to be 1.71, and the mean value of d' for R responses alone, d'_R, was found to be 1.80. Thus, unlike comparison of estimates of A', the mean value of d' based on R and K responses tended to *underestimate* the mean value based on R responses. In fact, d'_{RK} underestimated d'_R on 70% of the conditions examined by Donaldson (1996). However, this analysis was based on a relatively small sample of 17 studies and 80 experimental conditions. If it is extended to the current database of 400 conditions, then the mean values of d'_{RK} and d'_R are 1.520 and 1.527, respectively,[4] with d'_{RK} underestimating d'_R on only 53.7% of the conditions. Figure 2B presents the corresponding scatter plot of d'_{RK} against d'_R. Although some variability is present, the data points do not appear to depart systematically from the line of equality.

In summary, the proposal that estimates of sensitivity should be the same whether based on old–new decisions or RK decisions is confirmed as comparisons based on d' reveal little or no systematic differences. Although small but reliable differences are found for comparisons based on A', this can be attributed to the fact that A' varies systematically with the placement of the decision criterion if the underlying cumulative distribution functions are normal.

Remember, Know, and Guess Responses

The above argument concerning comparison of different estimates of sensitivity has also been examined with respect to a recent extension of the RK paradigm that includes an additional response category, corresponding to guesses (Gardiner, Ramponi, & Richardson-Klavehn, 2002). In the extended remember–know–

[4] In 13 conditions, the remember false alarm rate was zero. In order to calculate d', these were set equal to 0.005.

guess (RKG) paradigm, participants are instructed to respond "guess" if the target item does not fulfill the criteria for either an R or K response but it is judged to be old, perhaps on other strategic grounds. One of the reasons for including this category is to overcome a potential limitation of the original RK paradigm. By having only two response alternatives for recognized items, participants could find themselves in an ambiguous situation if they encounter an item that they suspect was presented in the study phase but that does not meet the criteria for either an R or K response. They are therefore faced with the choice of either responding "new," which would reduce their overall level of performance, or disobeying the instructions and either responding "know" or, less plausibly, "remember." Because participants are often motivated to perform well, it is possible that they may tend to choose the latter option. By allowing an additional response category for guesses, this problem is potentially avoided.

The SDT interpretation is readily extended to account for the RKG paradigm by the addition of two more functions to Equation 1. Let G_O be the probability of a guess response to an old item and let G_N be the probability of a guess response to a new item. Then,

$$R_O + K_O + G_O = 1 - F_O(g - d)$$

$$R_N + K_N + G_N = 1 - F_N(g),$$

where g is the decision criterion associated with guesses. It is assumed that $g \leq k \leq r$, in which case g now defines the old–new criterion. The equal-variance normal distribution model may be similarly extended.

Gardiner, Ramponi, and Richardson-Klavehn (2002) compared estimates of sensitivity based on R responses, R and K responses, and R, K, and G (guess) responses, with respect to a database of 23 RKG studies that comprised a total of 86 different experimental conditions. Across these conditions, they compared A'_R, A'_{RK}, and A'_{RKG}, the mean values of A' for R responses, for R and K responses, and for R, K, and G responses, respectively. According to the data in their Table 1, the corresponding values of A' are 0.786, 0.809, and 0.794, respectively,[5] and although the differences between these values are small, Gardiner et al. showed that they demonstrate a systematic trend. The estimates of A'_{RK} exceeded those of A'_R in 72% of conditions, they exceeded those of A'_{RKG} in 86% of conditions, and the estimates of A'_{RKG} exceeded those of A'_R in 59% of conditions. However, as discussed earlier, the expected value of A' depends on both discriminability and placement of the decision criterion, c. As Figure 3 shows, the function relating A' to c is curvilinear, reaching a maximum at intermediate values. The mean values of A' reported above suggest that the peak of this function should lie in the vicinity of the RK criterion because A'_{RK} is greater than either A'_R or A'_{RKG}. However, if the equal-variance normal distribution model is fitted to the data, the function peak is closer to the old–new criterion. For this model, the predicted mean values of A'_R, A'_{RK}, and A'_{RKG} are 0.792, 0.813, and 0.816, respectively, which appear to overestimate the observed means. Although the differences are small, they may reflect a real effect. This assumption is further supported by the equivalent analysis based on d'. In this case, the observed mean values of d'_R, d'_{RK}, and d'_{RKG} are 1.44, 1.40, and 1.25, respectively. Although the estimate of d' based on R responses is similar to the estimate based on R and K responses, both of these appear to be greater than the estimate based on R, K, and G responses. Although the precise implications of these data are unclear, it could suggest that the equal-variance signal detection model does not fit data from the RKG paradigm as well as it fits data from the RK paradigm. Yet, even so, this does not count necessarily as evidence against the SDT interpretation. It may only indicate that the assumptions underlying the equal-variance signal detection model have not been met. Specifically, if the distributions of old and new items are not normal or if their variances are not identical, then systematic departures from the predictions of the equal-variance signal detection model similar to those obtained above may be observed.

Argument 3: Predicting False Alarm Rates

In a recent study, Dobbins et al. (2000) proposed that the SDT interpretation could be refuted on the basis of statistical relationships observed between the old–new hit rate ($R_O + K_O$), the old–new false alarm rate ($R_N + K_N$), and the remember rate (R_O). They argued that under the SDT interpretation three specific relationships should be found:

> First, hit and false alarm rates will be positively correlated via the old/new response criterion. Second, remember and false alarm rates will be positively correlated because of the relationship between the remember[/know] and old/new response criteria. Finally, because the relation between the remember and false alarm rates is entirely redundant to that between the hit and false alarm rates, one would expect multiple regression to demonstrate that correct remember rates will yield no unique predictive information about false alarm rates. (Dobbins et al., 2000, p. 1349)

In order to test their three predictions, Dobbins et al. (2000) examined the distribution of R and K responses across 72 individuals who participated in three unrelated studies. Consistent with the first prediction, they found a positive correlation between the old–new hit rate and the old–new false alarm rate as well as between the old–new hit rate and the remember rate. However, in conflict with the second prediction, they found a nonsignificant *negative* correlation between the remember rate and the old–new false alarm rate. Furthermore, in conflict with the third prediction, they found a significantly stronger relationship between the old–new false alarm rate and the combination of the old–new hit rate and the remember rate than that between the old–new false alarm rate and the old–new hit rate alone. In addition, the remember rate added additional, *negative* information to the prediction of the old–new false alarm rate. This result was interpreted by Dobbins et al. in terms of a traditional suppression effect. They suggested that it occurred because "the remember rate shared a relationship with the hit rate that was independent of the false alarm behaviour" (Dobbins et al., 2000, p. 1352), and concluded that "the suppression effect of remember rates poses a problem for the signal detection account because it demonstrates that the hit rate has at least *two* unique and systematic sources of variance—one resulting from the relationship between remember and hit rates, the other between the remaining portion of hit rates (after remember rate is

[5] These are slightly different from those reported by Gardiner et al. (2002) because of errors in the calculation of the combined totals, $R + K$ and $R + K + G$, for Studies 10 and 11. The statistics reported in the present article are based on recalculation of these totals from the reported values of R, K, and G responses.

subtracted) and false alarm rates" (Dobbins et al., 2000, p. 1352, italics in original). Thus, according to Dobbins et al., these results are inconsistent with the SDT interpretation of the RK paradigm.

Notwithstanding the conclusions reached by Dobbins et al., the data they report are completely consistent with the SDT interpretation as the observed correlations are exactly as would be predicted by it. The predictions derived by Dobbins et al. are mistaken as they are based on an incomplete analysis of the SDT interpretation. In order to see this, it is necessary to reexamine the SDT functions for the three variables selected by Dobbins et al. From Equation 1, the expressions for these variables are as follows:

$$\text{old–new hit rate } (R_O + K_O) = 1 - F_O(k - d)$$

$$\text{old–new false alarm rate } (R_N + K_N) = 1 - F_N(k)$$

$$\text{remember rate } (R_O) = 1 - F_O(r - d). \tag{4}$$

The correlation between any pair of variables in Equation 4 depends on two principal factors. The first is the extent to which the component cumulative distribution functions, F_O and F_N, introduce significant nonlinear distortions. However, although this factor may tend to reduce the absolute value of any correlation, it is unlikely to introduce a correlation where none exists. The more important factor, as recognized by Dobbins et al. (2000), is the correlational structure of the component parameters. When variables share parameters in common, some of this correlation can be predicted directly from their equations (because any parameter is correlated perfectly with itself). But this is not the whole story. The overall correlation will also depend on the nature of the correlation between the remaining parameters that appear in the equation. This is an empirical question that will depend on properties of the data in question.

Consider the first prediction proposed by Dobbins et al. (2000), that the old–new hit rate should be positively correlated with the old–new false alarm rate. This prediction was based on the observation that both of these variables depend (negatively) on the old–new criterion, *k.* Yet examination of Equation 4 reveals that the overall hit rate depends on *d* as well as *k.* If the nonlinear effects of F_O and F_N are ignored, then the prediction of a positive correlation between the old–new hit rate and the old–new false alarm rate follows only if the parameters *k* and *d* are themselves not positively correlated. To the extent that they are positively correlated, the correlation between the two variables should decrease. Similar reasoning applies to the second prediction proposed by Dobbins et al., that there should be a positive correlation between the remember rate and the old–new false alarm rate. This prediction was based on a presumed positive correlation between *k* and *r* across participants. As stated by Dobbins et al. (2000), "The constraint that each subject must place his or her remember criterion higher on the strength continuum than the old/new criterion will mean that on average the two criteria . . . will be correlated across subjects" (p. 1349). Although this may be true, it ignores the possibility of a correlation between parameters *r* and *d.* As in the first prediction, the extent to which these parameters are themselves correlated will also affect the correlation between the old–new false alarm rate and the remember rate.

Are there reasons for supposing that the placement of a decision criterion, whether *k* or *r,* should be correlated with sensitivity, *d*? In relevant investigations of this question using different methodologies, Balakrishnan and Ratcliff (1996), Hirshman (1995), and Stretch and Wixted (1998) have all reported evidence that participants in recognition memory tasks shift their decision criteria in response to changes in discriminability. That is, in order to make approximately equal use of the response categories available to them, participants tend to adjust their decision criteria in order to partition the strength-of-evidence continuum into approximately equal parts. The result is a positive correlation between the decision criterion and the level of discriminability. Such behavior is also consistent with that of an unbiased observer who is attempting to maximize his or her overall level of performance while weighing the costs of misses and false alarms equally. Under the equal-variance signal detection model, such an unbiased observer should set *k* equal to $d'/2$. Thus, to the extent to which participants operate in this manner, the decision criteria, *k* and *r,* should be positively correlated with discriminability, *d.*

In their study, Dobbins et al. examined data from individual participants. However, an equivalent analysis can also be conducted on group data. Figure 4 presents data from the 400 conditions of the observed database. Figure 4A is a scatter plot of the relationship between the old–new hit rate and the old–new false alarm rate. In this case, there is a weak negative relationship between the two variables (coefficient $= -0.17$). In contrast, Dobbins et al. (2000) found a positive relationship between these variables. I return to this point below. First, however, it is necessary to point out that a weak negative relationship is not inconsistent with the SDT interpretation because it is also found in the model-specific database. Figure 5 presents the results of the same analyses conducted on the 400 conditions of the model-specific database. Figure 5A is a scatter plot of the old–new hit rate against the old–new false alarm rate. It is apparent that the same weak negative correlation is obtained (coefficient $= -0.16$). According to Equation 4, this suggests that there must be a positive correlation between *k* and *d* across the conditions. In the model-specific database, this correlation is 0.74. The positive correlation between the old–new hit rate and the old–new false alarm rate found by Dobbins et al. can therefore be explained in terms of a much weaker relationship between *k* and *d* in their data set. One reason for this may have been the fact that they based their analysis on comparisons of individuals under largely similar conditions rather than on comparisons of the means of groups of individuals under conditions designed to be as different as possible. It is probable that discriminability varies to a much greater extent between experimental conditions designed to manipulate this parameter than between participants tested under broadly similar conditions.

Figure 4B is a scatter plot of the relationship between the old–new hit rate and the remember rate in the observed database and shows, as also found by Dobbins et al. (2000), that there is a strong positive correlation between these variables (coefficient = 0.87). This correlation follows both from the common dependence of these variables on *d* as well as from the existence of positive correlations between *k* and *d* and between *k* and *r.* Figure 5B shows that the same relationship is observed in the model-specific database. For these data, the correlation between *k* and *r* is 0.83 and the resulting correlation between old–new hit rate and remember rate is 0.88, in close agreement with the observed correlation.

As their second prediction, Dobbins et al. proposed that there should be a positive correlation between the remember rate and the

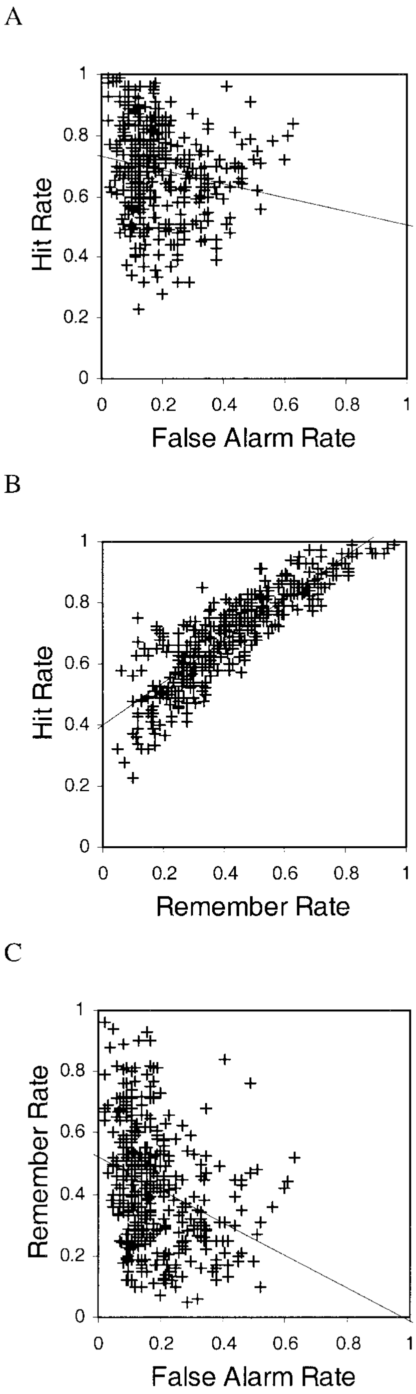


Figure 4. Scatter plots of the observed data. A: Relationship between the old–new hit rate and the old–new false alarm rate. B: Relationship between the old–new hit rate and the remember rate. C: Relationship between the remember rate and the old–new false alarm rate.

old–new false alarm rate. In their data however, they found a weak negative correlation of −0.18. A similar result occurs in the observed database. Figure 4C is a scatter plot of the remember rate against the old–new false alarm rate for each condition and shows a moderate negative relationship between these variables (coefficient = −0.32). Figure 5C is a scatter plot of the same variables from the model-specific database. Because the same negative correlation is found (coefficient = −0.32), it is clear that such a result is not inconsistent with an SDT interpretation. The reason for this correlation is that because *r* is necessarily less strongly

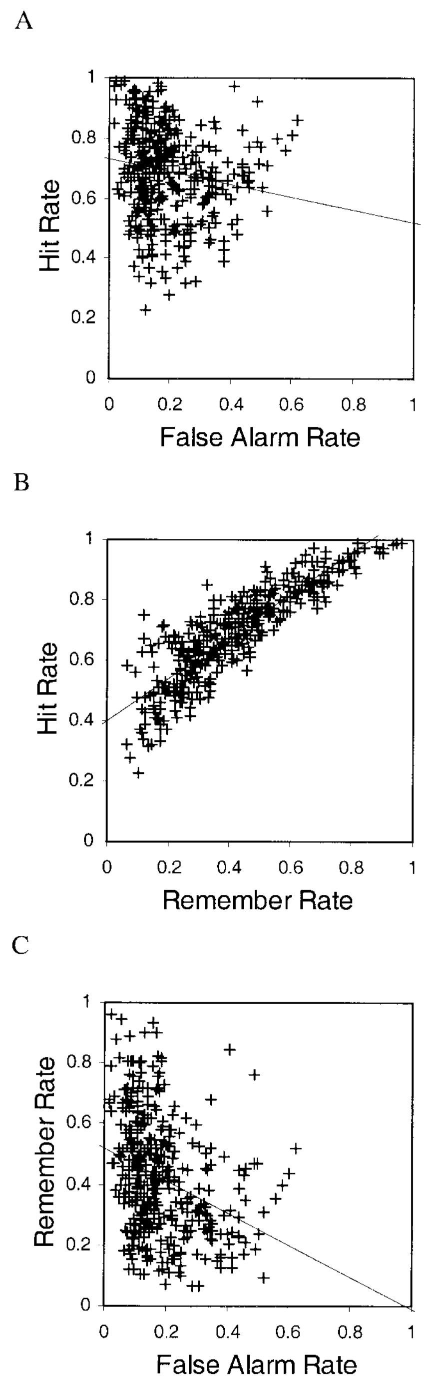


Figure 5. Scatter plots of model-specific data. A: Relationship between the old–new hit rate and the old–new false alarm rate. B: Relationship between the old–new hit rate and the remember rate. C: Relationship between the remember rate and the old–new false alarm rate.

correlated with k than is k itself, the correlation between the remember rate and the old–new false alarm rate is dominated to a greater extent by the (negative) contribution of the correlation between k and d.

Finally, consider the third prediction proposed by Dobbins et al. (2000), that in a multiple regression analysis, the remember rate should add no additional information to the prediction of the old–new false alarm rate than does the old–new hit rate alone. Regression analysis of the 400 conditions of the observed database yielded similar results to those found by Dobbins et al. That is, prediction of the false alarm rate is significantly improved if both the old–new hit rate and the remember rate are used as predictors than if the old–new hit rate alone is used ($\Delta R^2 = 0.13$, $p < 0.00001$). In addition, as also found by Dobbins et al., the contribution of the remember rate to the combined regression equation is negative. Yet contrary to the suggestion by Dobbins et al., this outcome is not inconsistent with SDT interpretation because the model-specific database yields exactly the same pattern. The reason for this finding is straightforward. Although the relationship between the old–new hit and false alarm rates is led to be positive by the fact that they share a common parameter, k, it is led to be negative (in other words, suppressed) to the extent that the parameters k and d are positively correlated. By adding the remember rate to the regression equation, and by making its contribution negative, the suppressing effect of d is removed. The result is a net improvement in prediction because r is less strongly correlated with k than is k itself.

Ironically, the analysis conducted by Dobbins et al. (2000) on their data, although presented as an argument against the SDT interpretation, is completely consistent with this account. They proposed that the old–new hit rate "has at least *two* unique and systematic sources of variance—one resulting from the relationship between remember and hit rates, the other between the remaining portion of hit rates (after remember rate is subtracted) and false alarm rates" (Dobbins et al., 2000, p. 1352). Examination of Equation 4 reveals that this follows directly from the SDT interpretation. First, the remember rate and the old–new hit rate are positively correlated primarily because they share a common parameter, d. If the remember rate is subtracted from the hit rate, the effect of this parameter is removed, leaving a difference between r and k. That is, if the effect of the nonlinear function, F_O, and different regression weights are ignored, we have

$$\begin{aligned}\text{old–new hit rate} - \text{remember rate} &\approx -(k-d) + (r-d)\\ &\approx r-k.\end{aligned}$$

If the old–new false alarm rate is written in the same way, ignoring the effect of the function, F_N, then,

$$\text{old–new false alarm rate} \approx -k.$$

Because r and k are not perfectly correlated, there is a residual correlation between the old–new false alarm rate and the difference between the old–new hit rate and the remember rate that results from the dependence of both these quantities on the shared parameter, k. Across the 400 conditions of the observed database, this correlation is a moderate 0.39.

In conclusion, because the predictions derived by Dobbins et al. (2000) are based on an incomplete analysis of the SDT interpretation, none of the effects that have been demonstrated are inconsistent with the SDT interpretation of the RK paradigm.

Argument 4: Confidence Judgments

The principal claim of the SDT interpretation is that RK responses are equivalent to confidence judgments. However, it has been argued that this is ruled out by the data because in studies in which both kinds of judgments are directly compared, different patterns of results are found (Gardiner, 2001; Parkin & Walter, 1992; Mäntylä, 1997; Rajaram, 1993; Rajaram, Hamilton, & Bolton, 2002). Data from two such studies, by Gardiner and Java (1990) and by Rajaram (1993), illustrate this point. Table 2 shows the results found by Gardiner and Java (1990). In one experiment of this study, participants were presented with either words or nonwords during the study phase and instructed to respond either "remember" or "know" to items that they recognized at test. A second experiment was identical to the first except that a different group of participants were instructed to respond either "sure" or "unsure" to items that they recognized at test. The results showed that for RK judgments, more R responses were observed for words than for nonwords, whereas the reverse occurred for K responses. In contrast, for sure–unsure judgments, whether the item was a word or nonword had almost no differential effect.

The results found by Rajaram (1993) are shown in Table 3. In this study, presentation of a word in the test phase was preceded by the brief masked presentation of either the same or an unrelated word. In one experiment, participants were asked to make RK

Table 2
Observed and Predicted Mean Proportions of Remember–Know and Sure–Unsure Responses from Gardiner and Java (1990)

Response category	Observed				Predicted			
	Old		New		Old		New	
	Words	Nonwords	Words	Nonwords	Words	Nonwords	Words	Nonwords
Remember	.28	.19	.04	.03	.26	.19	.08	.03
Know	.16	.30	.11	.12	.17	.30	.09	.12
Sure	.33	.39	.13	.07	.34	.39	.12	.09
Unsure	.28	.28	.22	.22	.29	.30	.21	.20

Table 3
Observed and Predicted Mean Proportions of Remember–Know and Sure–Unsure Responses from Rajaram (1993)

Response category	Observed				Predicted			
	Old		New		Old		New	
	Same	Different	Same	Different	Same	Different	Same	Different
Remember	.43	.42	.05	.05	.43	.42	.07	.06
Know	.24	.18	.18	.13	.26	.20	.14	.09
Sure	.57	.54	.06	.04	.55	.52	.12	.10
Unsure	.18	.16	.19	.16	.21	.18	.15	.11

judgments of recognized words. In a second otherwise identical experiment, participants were asked to make sure–unsure judgments of recognized words. The results showed that prior masked presentation of the same word increased K responses for both old and new words but had no effect on R responses. In contrast, both sure and unsure responses were increased by prior masked presentation of the same word, although the effect was significant only for new items. Thus, sure and unsure responses were similar to K responses but different from R responses.

Although the fact that experimental variables affect RK and sure–unsure judgments differently invites the inference that they are based on qualitatively different sources of information, this pattern of results is not inconsistent with the SDT interpretation. Tables 2 and 3 also show the predicted mean proportions of responses derived from the equal-variance normal distribution model. The model was fitted separately to the data from each study under the constraint that the estimate of d' for each level of the experimental variable condition should be the same under each set of instructions (RK vs. sure–unsure).[6] It is apparent that the model fits the data well and captures the observed differences between RK and sure–unsure instructions in both studies. That is, the effects of the experimental variables on each response category are the same for the predicted means as they are for the observed means.

The best-fitting model parameters are shown in Table 4. For the study by Gardiner and Java (1990), the estimate of d' is greater for nonwords than for words. This finding is consistent with the well-replicated effect of word frequency on recognition memory (Schulman, 1967). Of greater interest is the effect of instructions on the values of the decision criteria. Under RK instructions, participants appear to adopt relatively stringent criteria whereas under sure–unsure instructions, both criteria are more lenient. Under RK instructions, participants were told that, “often, when remembering a previous event or occurrence, we consciously recollect and become aware of aspects of the previous experience. At other times, we simply know that something has occurred before, but without being able consciously to recollect anything about its occurrence or what we experienced at the time” (Gardiner & Java, 1990, p. 25). That is, for an item to be recognized, it must be either “remembered” or “known” in this sense. In contrast, under sure–unsure instructions, participants were simply asked to classify each recognized item as either “sure” or “unsure.” It seems reasonable that participants would, under these circumstances, adopt a more lenient set of criteria. Of interest is that the same effect is also found in the data from Rajaram (1993). Although in this case there is little or no effect of prime type on d', the relative placement of the decision criteria is more lenient under sure–unsure instructions than under RK instructions.

Table 4
Best-Fitting Parameters of the Equal-Variance Normal Distribution Model for Each Condition of Gardiner and Java (1990) and Rajaram (1993)

Model parameter	Gardiner & Java (1990)		Rajaram (1993)	
	Words	Nonwords	Same	Different
Sensitivity (d')	0.79	1.04	1.32	1.33
Criteria (c)				
Remember	1.42	1.91	1.49	1.52
Know	0.95	1.06	0.81	1.02
Sure	1.19	1.32	1.20	1.28
Unsure	0.45	0.54	0.63	0.80

The results of a study by Mäntylä (1997) provide more detailed information concerning the relationship between confidence ratings and RK decisions. In one experiment, participants classified recognized items into one of three categories, remember or R responses, know or K responses, and guess or G responses. In a second experiment, identical in all other respects to the first, participants were asked to classify recognized items on a 4-point confidence scale labeled *very high, high, low*, and *very low*, respectively. The stimuli were photographs of faces that participants rated in terms of their distinctive or relational characteristics during the study phases of both experiments. Under RKG instructions, there was a significant effect of encoding task in one direction for R responses and in the other direction for K responses. In contrast, there was little or no effect of this variable on any of the categories of the confidence scale. Mäntylä (1997) concluded that these results, although consistent with the dual-process interpretation, were inconsistent with the SDT interpretation.

Although Mäntylä (1997) found a significant difference between the different encoding instructions for both R and K responses, a reanalysis of these data reveals that they are not inconsistent with the SDT interpretation. Table 5 presents the observed

[6] Even though small differences in d' are likely to occur between different groups of subjects.

Table 5
Observed and Predicted Data and Best-Fitting Parameters of the Equal-Variance Normal Distribution Model for Experiments 1 and 2 of Mäntylä (1997)

Response category	Observed[a]			Predicted			d'		Criterion c
	Dist.	Rel.	New	Dist.	Rel.	New	Dist.	Rel.	
Experiment 1							2.27	2.03	
Remember	0.54	0.45	0.03	0.55	0.45	0.02			2.15
Know	0.20	0.28	0.04	0.23	0.25	0.05			1.51
Guess	0.11	0.13	0.09	0.11	0.14	0.08			1.05
Experiment 2							2.10	2.05	
Very high	0.45	0.43	0.01	0.45	0.43	0.01			2.22
High	0.20	0.20	0.03	0.20	0.20	0.03			1.72
Low	0.12	0.12	0.05	0.12	0.12	0.04			1.36
Very low	0.07	0.08	0.04	0.07	0.07	0.05			1.12

Note. Dist. = distinctive condition; Rel. = relational condition; d' = sensitivity.
[a] Observed data from Experiment 1 have been estimated from Figure 1 of Mäntylä (1997, p. 1207).

mean proportion of responses in each of the categories used in Experiment 1 (RKG instructions) and Experiment 2 (confidence ratings). Also shown are the predicted means based on the equal-variance normal distribution model. This model was fitted separately to each experiment under the constraint that the decision criteria should be the same under both distinctive and relational conditions. Three aspects of these data merit consideration. First, although there are small differences in the estimates of sensitivity (d') between the two experiments, they are similar both in magnitude and in demonstrating a small advantage for the distinctive condition over the relational condition. Second, these relatively minor differences, coupled with different placements of the decision criteria, are sufficient to capture the observed difference in R responses between the distinctive and relational conditions as well as, albeit to a lesser extent, the difference in K responses. This is in contrast to the similarity in confidence ratings between the two conditions. Third, comparison of the two sets of decision criteria provides some information concerning how participants may have interpreted RKG instructions. Specifically, R responses appear to correspond to a level of confidence between "high" and "very high," K responses appear to correspond to a level of confidence between "low" and "high," and G responses appear to correspond to a level of confidence between "very low" and "low."

Argument 5: Dissociation of Remember and Overall Hit Rates

In a recent article, Conway et al. (2001) proposed that the SDT interpretation is refuted if the proportion of R responses is found to differ between two conditions while the overall hit rate remains the same. In a series of experiments, they compared two different encoding instructions that were either relatively high or low in self-reference. In the first experiment of the series, Conway et al. observed a higher proportion of R responses and a lower proportion of K responses to targets in the high self-reference condition than to targets in the low self-reference condition. Because there was an increase in R responses and a decrease in K responses, the net effect was no difference in the overall hit rate. There was thus a dissociation between remember rate and overall hit rate from which the authors concluded that "these differences cannot be explained by a trace-strength account of memory" (Conway et al., 2001, p. 680).

The argument advanced by Conway et al. is not correct. Let R_1 and R_2 be the remember rates for two conditions and let H_1 and H_2 be the corresponding overall hit rates. Then, from Equation 1, the difference in remember rates is given by

$$R_1 - R_2 = F_O(r_2 - d_2) - F_O(r_1 - d_1),$$

and the difference in overall hit rates is given by

$$H_1 - H_2 = F_O(k_2 - d_2) - F_O(k_1 - d_1).$$

Clearly, the fact that the second of these equations is equal to zero does not imply that the first equation should also be equal to zero.

Explanatory Power of the SDT Interpretation

Each of the arguments evaluated in the previous sections has been shown to be flawed. Hence, in relation to these arguments, there currently exists no data from either the RK or RKG paradigms that can be shown to be inconsistent with the SDT interpretation. Yet it is not sufficient that a viable model be merely consistent with the data, it should also offer a plausible explanation of the phenomena of interest. For the SDT interpretation, this means that the parameters *d*, *k*, and *r* ought to be affected by appropriate experimental manipulations in psychologically meaningful ways. It is in relation to this point that Gardiner and Richardson-Klavehn (2000) have criticized the SDT interpretation, suggesting that it fails "to explain why the criteria [model parameters] are affected by different independent and subject variables in the ways they have to be to fit the data" (p. 236). This raises an important issue that has yet to be systematically investigated in the literature, partly as a consequence of the view that the SDT interpretation has already been ruled out by the data. Although it is beyond the scope of the present article to discuss in detail whether or not the parameters of the SDT interpretation are affected systematically by all possible experimental factors, I will attempt to illustrate its explanatory power in relation to the effects of three different factors on R and K responses. These factors concern the proportion of old words presented at test, the level of

Table 6
Fit of the Equal-Variance Normal Distribution Model to the Results of Studies Varying the Proportion of Old Items Presented at Test

Study and condition	Observed Old R	Observed Old K	Observed New R	Observed New K	Best-fitting model parameters d'	k	r	Predicted Old R	Predicted Old K	Predicted New R	Predicted New K
Strack & Förster (1995)											
30% old	0.30	0.26	0.03	0.10	1.23	1.11	1.79	0.30	0.26	0.04	0.10
50% old	0.37	0.33	0.02	0.23	1.23	0.67	1.59	0.37	0.35	0.06	0.19
Hirshman & Henzler (1998)											
30% old	0.19	0.29	0.05	0.13	0.74	0.84	1.64	0.19	0.28	0.05	0.15
70% old	0.28	0.43	0.11	0.35	0.74	0.15	1.29	0.29	0.43	0.10	0.34
Verfaellie et al. (2001)											
Amnesic											
30% old	0.10	0.38	0.03	0.22	0.82	0.79	2.03	0.11	0.40	0.02	0.19
70% old	0.22	0.44	0.07	0.21	0.82	0.53	1.64	0.21	0.41	0.05	0.25
Controls											
30% old	0.50	0.18	0.01	0.04	2.00	1.54	2.00	0.50	0.18	0.02	0.04
70% old	0.60	0.19	0.03	0.10	2.00	1.14	1.73	0.60	0.20	0.04	0.09

Note. R = mean proportion of remember responses; K = mean proportion of know responses; d' = sensitivity; k = old–new decision criterion; r = remember–know decision criterion.

processing demands of the orienting task given at study, and comparison of amnesic participants and controls.

Three studies have investigated the effect of inducing a change in response criteria on the proportion of R and K responses (Strack & Förster, 1995; Hirshman & Henzler, 1998; Verfaellie, Giovanello, & Keane, 2001). In each of these studies, participants were told that either a relatively small or large proportion of studied words would appear in the test phase. In fact, half the test items were from the studied list in each case. Under these circumstances, according to SDT (Green & Swets, 1966), if participants are motivated to maximize their overall level of performance, they should adopt a more stringent criterion for responding "old" in the low-proportion condition than in the high-proportion condition. The results of these studies are shown in Table 6 along with the values of the best-fitting parameters and corresponding predictions of the equal-variance normal distribution model. This model was fitted separately to each pair of conditions under the constraint that d' should be the same under each level of the independent variable in each study. The model fits the data well and in each case a shift in decision criteria in the expected direction is observed. That is, when participants are told that a relatively high proportion of test items are from the studied list, the criteria decrease for both old–new and RK decisions. These results pose difficulties for the dual-process interpretation because according to this interpretation, R and K responses are assumed to reflect different forms of memory retrieval. If an experiment reports an increase in both kinds of response, as do the experiments in Table 6, this must be attributed to an increase in the success of the different forms of memory retrieval. However, it is difficult to see how instructions concerning the proportion of old items to be expected in a recognition test should have any effect on retrieval from memory.

Table 7 presents data from seven studies that have investigated the effect on the proportion of R and K responses of a change in the level of processing demanded by the orienting task given at study.[7] Some of these data have been averaged over other experimental factors in order to clarify the results. In each study, the false alarm rate was constrained to be equal across each level of the independent variable. Thus, in contrast to the studies listed in Table 6 that attempted to selectively manipulate the decision criteria while keeping sensitivity constant, the studies listed in Table 7 attempted to selectively manipulate sensitivity while keeping the decision criteria constant.

The results presented in Table 7 show that a change from shallow to deep encoding instructions consistently increases the proportion of R responses. According to the dual-process interpretation, this should be due to an improvement in the form of memory retrieval that underlies R responses. According to the SDT interpretation, it is attributable to an increase in sensitivity. Also presented in Table 7 are the values of the best-fitting parameters and corresponding predictions of the equal-variance normal distribution model. This model was fitted separately to each pair of conditions under the constraint that k and r should be the same under each level of the independent variable in each study. It is apparent that the difference in R responses is well accounted for by a concomitant change in d'.

Table 7 also demonstrates that level of processing manipulations have more variable effects on K responses. Whereas the original study by Gardiner (1988) found no change in K responses between shallow and deep processing at study, subsequent studies have found either an increase (two studies) or a decrease (four studies). This pattern is less clearly explained by the dual-process interpretation, at least if it is assumed that R and K responses directly reflect each underlying process. One dual-process model that departs from this view is the independence model proposed by Yonelinas and Jacoby (1995). In this model, R and K responses reflect the levels of two independent memory processes, called recollection and familiarity. Recollection is indexed by the pro-

[7] A wide variety of manipulations have been used. For the present analysis, "generate" versus "read" instructions have been classified as "deep" and "shallow," respectively.

Table 7
Fit of the Equal-Variance Normal Distribution Model to the Results of Studies Varying Levels of Processing Instructions

	Observed				Best-fitting model parameters			Predicted			
	Old		New					Old		New	
Study and condition	*R*	*K*	*R*	*K*	*d′*	*k*	*r*	*R*	*K*	*R*	*K*
Gardiner (1988, Exp. 1)											
Shallow	0.35	0.17	0.06	0.07	1.22	1.17	1.61	0.35	0.17	0.05	0.07
Deep	0.65	0.17	0.05	0.07	2.10	1.17	1.61	0.69	0.13	0.05	0.07
Khoe et al. (2000, Exp. 1)											
Shallow	0.16	0.29	0.08	0.25	0.32	0.43	1.34	0.15	0.30	0.09	0.24
Deep	0.26	0.34	0.08	0.25	0.70	0.43	1.34	0.26	0.35	0.09	0.24
Perfect et al. (1995, Exp. 2a)											
Shallow	0.37	0.28	0.02	0.05	2.13	1.76	2.46	0.37	0.27	0.01	0.03
Deep	0.69	0.19	0.02	0.01	2.95	1.76	2.46	0.69	0.19	0.01	0.03
Rajaram (1993, Exp. 1)											
Shallow	0.32	0.30	0.02	0.14	1.28	0.96	1.75	0.32	0.31	0.04	0.13
Deep	0.66	0.20	0.02	0.14	2.15	0.96	1.75	0.66	0.23	0.04	0.13
Schacter, Curran, et al. (1996, Exp. 1)											
Shallow	0.37	0.23	0.05	0.12	1.23	0.95	1.56	0.37	0.24	0.06	0.11
Deep	0.90	0.06	0.05	0.12	2.87	0.95	1.56	0.90	0.07	0.06	0.11
Wippich (1992)											
Shallow	0.17	0.24	0.02	0.12	0.82	1.05	1.79	0.17	0.25	0.04	0.11
Deep	0.40	0.31	0.02	0.12	1.62	1.05	1.79	0.43	0.28	0.04	0.11
Yonelinas (2001)											
Shallow	0.19	0.32	0.03	0.18	0.80	0.78	1.67	0.19	0.32	0.05	0.17
Deep	0.54	0.27	0.03	0.18	1.80	0.78	1.67	0.55	0.30	0.05	0.17

Note. R = mean proportion of remember responses; K = mean proportion of know responses; d' = sensitivity; k = old–new decision criterion; r = remember–know decision criterion; Exp. = experiment.

portion of R responses, and familiarity is indexed by the proportion of K responses divided by one minus the proportion of R responses (for a detailed discussion and derivation, see Yonelinas, 2002). If these formulas are applied to the data in Table 7, the variability in K responses disappears. For each study, a shift from shallow to deep processing is accompanied by an increase in *both* recollection and familiarity. This is important evidence in support of this model.

The variable effect of levels of processing on K responses may also be explained by the SDT interpretation. Table 7 shows that the proportions of K responses are reasonably well fit by the equal-variance normal distribution model, and, more important, that this model also captures the variable effect of level of processing on K responses. According to the SDT interpretation, for the data presented in Table 7, any differences between shallow and deep conditions are attributable solely to a change in sensitivity. For R responses, this always means that there is an increase from shallow to deep. For K responses, if the strength-of-evidence distribution of old items is unimodal then whether there is an increase or decrease depends crucially on the relationship between d' and the two decision criteria. Intuitively, if d' is very much less than k or very much greater than r, then the proportion of K responses for old items is at a minimum, but if d' falls between k and r, then the proportion of K responses is at a maximum. For the equal-variance normal distribution model, this maximum occurs when $d' = (k + r)/2$. Let d'_S and d'_D be the sensitivities associated with shallow and deep processing, respectively, and let $d'_{\max} = (k + r)/2$. If d'_S and d'_D are both less than $d'_{\max}$, then there should be an increase in the proportion of K responses from shallow to deep processing. If d'_S is less than $d'_{\max}$ and d'_D is greater than $d'_{\max}$ (by approximately the same amount), then there should be little or no change in the proportion of K responses. Finally, if d'_S and d'_D are both greater than $d'_{\max}$, then there should be a decrease in the proportion of K responses from shallow to deep processing. This pattern accounts for all of the data presented in Table 7.

The foregoing analysis also leads to the following prediction of the SDT interpretation. If the strength-of-evidence distribution of old items is unimodal and the false alarm rate for R and K responses is fixed across a set of conditions, then over a sufficient range of R responses, the relationship between the proportion of R responses and the proportion of K responses should be nonmonotonic. Specifically, as the proportion of R responses increases, the proportion of K responses should first increase and then decrease. This is a necessary consequence of the SDT interpretation (and the assumption of unimodality) and has nothing to do with the nature of the independent variable that affects d'. In contrast, although such a result is consistent with the dual-process independence model, it is not demanded by it. According to this model, if both familiarity and recollection increase across the levels of an independent variable, a similar nonmonotonic relationship between R and K responses will result. However, because familiarity and recollection are assumed to be independent, other relationships between R and K responses may occur, thereby refuting the SDT interpretation.

Table 8 presents summary data from four studies that compared amnesic participants and controls on the RK paradigm. As in Table 7, the data from some studies have been averaged over other independent variables in order to clarify the effect of amnesia on R and K responses. Although it has been suggested that amnesia has a selective effect on R responses (Hirshman, Fisher, Henthorn,

Table 8
Fit of the Equal-Variance Normal Distribution Model to the Results of Studies Comparing Groups of Amnesic Patients and Controls

	Observed				Best-fitting model parameters			Predicted			
	Old		New					Old		New	
Study and condition	*R*	*K*	*R*	*K*	*d′*	*k*	*r*	*R*	*K*	*R*	*K*
Knowlton & Squire (1995)											
Amnesic	0.17	0.16	0.09	0.10	0.45	0.89	1.39	0.18	0.15	0.08	0.10
Control	0.46	0.21	0.02	0.04	1.98	1.54	2.08	0.46	0.21	0.02	0.04
Schacter, Verfaellie, & Pradere (1996)											
Amnesic	0.28	0.23	0.14	0.20	0.45	0.43	1.05	0.27	0.24	0.15	0.19
Control	0.71	0.13	0.06	0.12	2.02	0.92	1.45	0.72	0.15	0.07	0.10
Schacter et al. (1997)											
Amnesic	0.18	0.29	0.07	0.14	0.75	0.84	1.65	0.18	0.28	0.05	0.15
Control	0.31	0.31	0.05	0.12	1.27	0.97	1.76	0.31	0.31	0.04	0.13
Verfaellie et al. (2001)											
Amnesic	0.16	0.41	0.05	0.22	0.83	0.66	1.82	0.16	0.41	0.03	0.22
Control	0.36	0.31	0.04	0.13	1.42	0.98	1.77	0.36	0.31	0.04	0.13

Note. R = mean proportion of remember responses; K = mean proportion of know responses; d' = sensitivity; k = old–new decision criterion; r = remember–know decision criterion.

Arndt, & Passannante, 2002), the results of these four studies present a mixed picture. Even though in each study the proportion of R responses was found to be less for amnesics than controls, changes in the proportion of K responses were more variable. One study found relatively little change in K responses; however, another found a decrease, and the remaining two found an increase.

As discussed in relation to the effects of levels of processing, the variability in K responses poses difficulties for some dual-process accounts. Nevertheless, it can be accommodated by the independence model if it is assumed that amnesia affects both recollection and familiarity. Yet the data may also be explained in terms of the SDT interpretation. The best-fitting parameter values and predictions of the equal-variance normal distribution model are also presented in Table 8. Although no claim is made that this is the correct SDT model, it can be seen that it is able to capture the main features of the data. More important, the model demonstrates several regularities. First, as expected in each study, there is a decrease in d' for the amnesic groups compared with the controls. Second, there is also a decrease in the levels of both decision criteria, k and r. This effect follows directly from the principle that the placement of the decision criterion is moderated by the level of sensitivity (Hirshman, 1995). Although the optimal placement of this criterion, under the equal-variance normal distribution model, is at $d'/2$, participants must also balance at least two other demands. First, they are required to follow, as far as possible, the RK instructions. As suggested by the results found by Mäntylä (1997), these instructions invite participants to adopt relatively stringent criteria for both old–new and RK decisions. As a result, participants would be encouraged to place the old–new criterion, k, at a point on the strength-of-evidence continuum that is numerically greater than optimal. Second, it is not unreasonable, given these instructions, to suppose that participants may expect that a proportion of responses should fall in both categories. In other words, they are likely to be reluctant to set the RK criterion, r, at such a stringent level that effectively no items are "remembered." It is conceivable that the net effect of these demands would be to position the decision criteria more or less as they appear to have been according to the data in Table 8. As a result, although the proportion of R responses (and the proportion of overall hits) will necessarily be less in amnesics compared with controls (because of the decrease in d'), the difference in the proportion of K responses will be more variable because of the different emphases that participants place on each of the competing demands that they are attempting to satisfy.

In conclusion, at least in relation to the examples given above, the effects of different variables on the parameters of the SDT interpretation appear to follow a number of regularities and to be explicable in terms of a few relatively simple considerations. In light of this, it should be possible to develop more detailed and psychologically meaningful accounts of the RK paradigm within an SDT framework.

Summary and Conclusions

The aim of the present article has been to critically evaluate the proposal that the SDT interpretation of the RK paradigm has been ruled out by the evidence. Although not constituting an exhaustive list, five prominent arguments against the SDT interpretation were examined. Each of these arguments proposes that the SDT interpretation is unable to account for a particular pattern of data that can be shown to exist. Yet in each case, it could be shown that the pattern of data is not inconsistent with the SDT interpretation. Therefore, this interpretation is not ruled out by these data. This is not to say that the SDT interpretation is consistent with all possible outcomes. Despite the fact that existing tests are flawed, it is possible to derive other patterns of data that are inconsistent with the SDT interpretation. In addition, although further examination of this question is required, it is possible to outline some general principles that govern the effect of different experimental variables on the theoretical constructs implied by the SDT interpretation. These principles concern the relationship between sensitivity and decision criteria and provide a framework for explaining observed changes in R and K responses.

The principal conclusion of the present study is that the SDT interpretation represents a psychologically plausible and empirically viable account of the RK paradigm and, by extension, of the RKG paradigm. This conclusion, however, requires two qualifications. First, it should not be inferred that the alternative dual-process interpretation has been shown to be incorrect. Rather, on the basis of the evidence reviewed in the present article, this interpretation has merely not yet been shown to be correct. Future research may yet reveal that the SDT interpretation is ruled out by the data. Second, the present conclusion applies only to data from the RK and RKG paradigms—it does not address other kinds of data that have been used to support dual-process accounts of recognition (e.g., Jacoby, 1991; Mandler, 1980; Yonelinas, 2002). It should also be noted that even if recognition memory can be decomposed into two qualitatively different components, it does not mean that these components are necessarily indexed by R and K judgments.

The conclusion reached in the present study is similar to the view proposed by Hirshman, Lanning, et al. (2002), who have urged "investigators of the remember–know paradigm and of memory consciousness to use the two-criterion signal detection model to facilitate theoretical development" (p. 155). By heeding this advice, it should be possible to determine whether the RK paradigm indexes two separate components of memory or whether it is simply a matter of confidence.

References

Balakrishnan, J. D., & Ratcliff, R. (1996). Testing models of decision making using confidence ratings in classification. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 615–633.

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

Conway, M. A., Dewhurst, S. A., Pearson, N., & Sapute, A. (2001). The self and recollection reconsidered: How a "failure to replicate" failed and why trace strength accounts of recollection are untenable. *Applied Cognitive Psychology, 15,* 673–686.

DeCarlo, L. T. (2002). Signal detection theory with finite mixture distributions: Theoretical developments with applications to recognition memory. *Psychological Review, 109,* 710–721.

Dobbins, I. G., Khoe, W., Yonelinas, A. P., & Kroll, N. E. A. (2000). Predicting individual false alarm rates and signal detection theory: A role for remembering. *Memory & Cognition, 29,* 1347–1356.

Donaldson, W. (1993). Accuracy of d' and A' as estimates of sensitivity. *Bulletin of the Psychonomic Society, 31,* 271–274.

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

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

Gardiner, J. M. (2000). On the objectivity of subjective experiences of autonoetic and noetic consciousness. In E. Tulving (Ed.), *Memory, consciousness, and the brain: The Talinn Conference* (pp. 159–172). Philadelphia: Psychology Press.

Gardiner, J. M. (2001). Episodic memory and autonoetic consciousness: A first-person approach. *Philosophical Transcripts of the Royal Society of London B, 356,* 1351–1361.

Gardiner, J. M., & Conway, M. A. (1999). Levels of awareness and varieties of experience. In B. M. Challis & B. M. Velichovsky (Eds.), *Stratification in cognition and consciousness* (pp. 237–254). Amsterdam: John Benjamins.

Gardiner, J. M., & Gregg, V. H. (1997). Recognition memory with little or no remembering: Implications for a detection model. *Psychonomic Bulletin & Review, 4,* 474–479.

Gardiner, J. M., & Java, R. I. (1990). Recollective experience in word and nonword recognition. *Memory & Cognition, 18,* 23–30.

Gardiner, J. M., Kaminska, Z., Dixon, M., & Java, R. L. (1996). Repetition of previously novel melodies sometimes increases both remember and know responses in recognition memory. *Psychonomic Bulletin & Review, 3,* 366–371.

Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (1998). Experiences of remembering, knowing, and guessing. *Consciousness and Cognition, 7,* 1–26.

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,* 83–98.

Gardiner, J. M., & Richardson-Klavehn, R. (2000). Remembering and knowing. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 229–244). Oxford, England: Oxford University Press.

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

Gregg, V. H., & Gardiner, J. H. (1994). Recognition memory and awareness: A large effect of study-test modalities on "know" responses following a highly perceptual orienting task. *European Journal of Cognitive Psychology, 6,* 131–147.

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin, 75,* 424–429.

Hirshman, E. (1995). Decision processes in recognition memory: Criterion shifts and the list-strength paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 302–313.

Hirshman, E. (1998). On the utility of the signal detection model of the remember–know paradigm. *Consciousness and Cognition, 7,* 103–107.

Hirshman, E., Fisher, J., Henthorn, T., Arndt, J., & Passannante, A. (2002). Midazolam amnesia and dual-process models of the word-frequency mirror effect. *Journal of Memory and Language, 47,* 499–516.

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

Hirshman, E., Lanning, K., Master, S., & Henzler, A. (2002). Signal-detection models as tools for interpreting judgments of recollections. *Applied Cognitive Psychology, 16,* 151–156.

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

Inoue, C., & Bellezza, F. S. (1998). The detection model of recognition using know and remember judgments. *Memory & Cognition, 26,* 299–308.

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

Jacoby, L. L., Yonelinas, A. P., & Jennings, J. M. (1997). The relation between conscious and unconscious (automatic) influences: A declaration of independence. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 13–47). Mahwah, NJ: Erlbaum.

Khoe, W., Kroll, N. E. A., Yonelinas, A. P., Dobbins, I. A., & Knight, R. T. (2000). The contribution of recollection and familiarity to yes–no and forced-choice recognition tests in healthy subjects and amnesics. *Neuropsychologia, 38,* 1333–1341.

Knowlton, B. J., & Squire, L. R. (1995). Remembering and knowing: Two different expressions of declarative memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 699–710.

Macmillan, N. A., & Creelman, C. D. (1996). Triangles in ROC space: History and theory of "nonparametric" measures of sensitivity and response bias. *Psychonomic Bulletin & Review, 3,* 164–170.

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

Mäntylä, T. (1997). Recollection of faces: Remembering differences and knowing similarities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1203–1216.

Murdock, B. B. (1974). *Human memory: Theory and data.* Potomac, MD: Erlbaum.

Parkin, A. J., & Walter, B. M. (1992). Recollective experience, normal aging and frontal dysfunction. *Psychology and Aging, 7,* 290–298.

Perfect, T. J., Williams, R. B., & Anderton-Brown, C. (1995). Age differences in reported recollective experience are due to encoding effects, not response bias. *Memory, 3,* 169–186.

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

Rajaram, S. (1996). Perceptual effects on remembering: Recollective processes in picture recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 365–377.

Rajaram, S., Hamilton, M., & Bolton, A. (2002). Distinguishing states of awareness from confidence during retrieval: Evidence from amnesia. *Cognitive, Affective, & Behavioral Neuroscience, 2,* 227–235.

Schacter, D. L., Curran, T., Galluccio, L., Milberg, W. P., & Bates, J. F. (1996). False recognition and the right frontal lobe: A case study. *Neuropsychologia, 34,* 793–808.

Schacter, D. L., Verfaellie, M., & Anes, M. D. (1997). Illusory memories in amnesic patients: Conceptual and perceptual false recognition. *Neuropsychology, 11,* 331–342.

Schacter, D. L., Verfaellie, M., & Pradere, D. (1996). The neuropsychology of memory illusions: False recall and recognition in amnesic patients. *Journal of Memory and Language, 35,* 319–334.

Schulman, A. I. (1967). Word length and rarity in recognition memory. *Psychonomic Science, 9,* 211–212.

Strack, F., & Förster, J. (1995). Reporting recollective experiences: Direct access to memory systems? *Psychological Science, 6,* 352–358.

Stretch, V., & Wixted, J. T. (1998). Decision rules for recognition memory confidence judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1397–1410.

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

Verfaellie, M., Giovanello, K. S., & Keane, M. M. (2001). Recognition memory in amnesia: Effects of relaxing response criteria. *Cognitive, Affective, & Behavioral Neuroscience, 1,* 3–9.

Wippich, W. (1992). Implicit and explicit memory without awareness. *Psychological Research, 54,* 212–224.

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

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.

Appendix

Studies Included in the Observed Database

Blaxton, T. A., & Theodore, W. H. (1997). The role of the temporal lobes in recognizing visuospatial materials: Remembering versus knowing. *Brain and Cognition, 35,* 5–25.

Bowler, D. M., Gardiner, J. M., & Grice, S. (2000). Episodic memory and remembering in adults with Asperger syndrome. *Journal of Autism and Developmental Disorders, 30,* 295–304.

Bowler, D. M., Gardiner, J. M., Grice, S., & Saavalainen, P. (2000). Memory illusions: False recall and recognition in adults with Asperger's syndrome. *Journal of Abnormal Psychology, 109,* 663–672.

Clarys, D., Insingrini, M., & Kamel, G. (2002). Mediators of age-related differences in recollective experience in recognition memory. *Acta Psychologica, 109,* 315–329.

Conway, M. A., & Dewhurst, S. A. (1995a). Remembering, familiarity and source monitoring. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 48*(A), 125–140.

Conway, M. A., & Dewhurst, S. A. (1995b). The self and recollective experience. *Applied Cognitive Psychology, 9,* 1–19.

Curran, H. V., Gardiner, J. M., Java, R. I., & Allen, D. (1993). Effects of lorazepam upon recollective experience in recognition memory. *Psychopharmacology, 110,* 374–378.

Danion, J.-M., Rizzo, L., & Bruant, A. (1999). Functional mechanisms underlying impaired recognition memory and conscious awareness in patients with schizophrenia. *Archives of General Psychiatry, 56,* 639–644.

Dewhurst, S. A., & Anderson, S. J. (1999). Effects of exact and category repetition in true and false recognition memory. *Memory & Cognition, 27,* 665–673.

Dewhurst, S. A., & Conway, M. A. (1994). Pictures, images, and recollective experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1088–1098.

Dewhurst, S. A., & Hitch, G. J. (1997). Illusions of familiarity caused by cohort activation. *Psychonomic Bulletin & Review, 4,* 566–571.

Dewhurst, S. A., Hitch, G. J., & Barry, C. (1998). Separate effects of word frequency and age of acquisition in recognition and recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 284–298.

Dobbins, I. G., Kroll, N. E. A., & Liu, Q. (1998). Confidence–accuracy inversions in scene recognition: A remember–know analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1306–1315.

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

Eldridge, L. L., Sarfatti, S., & Knowlton, B. J. (2002). The effect of testing procedure on remember–know judgments. *Psychonomic Bulletin & Review, 9,* 139–145.

Gallo, D. A., Roberts, M. J., & Seamon, J. G. (1997). Remembering words not presented in lists: Can we avoid creating false memories? *Psychonomic Bulletin & Review, 4,* 271–276.

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

Gardiner, J. M., Gawlik, B., & Richardson-Klavehn, A. (1994). Maintenance rehearsal affects knowing, not remembering; elaborative rehearsal affects remembering, not knowing. *Psychonomic Bulletin & Review, 1,* 107–110.

Gardiner, J. M., & Gregg, V. H. (1997). Recognition memory with little or no remembering: Implications for a detection model. *Psychonomic Bulletin & Review, 4,* 474–479. (Includes unpublished data from Gregg & Gardiner, 1994)

Gardiner, J., Gregg, V. H., Mashru, R., & Thaman, M. (2001). Impact of encoding depth on awareness of perceptual effects in recognition memory. *Memory & Cognition, 29,* 433–440.

Gardiner, J. M., & Java, R. I. (1990). Recollective experience in word and nonword recognition. *Memory & Cognition, 18,* 23–30.
Gardiner, J. M., & Java, R. I. (1991). Forgetting in recognition memory with and without recollective experience. *Memory & Cognition, 19,* 617–623.
Gardiner, J. M., Kaminska, Z., Dixon, M., & Java, R. L. (1996). Repetition of previously novel melodies sometimes increases both remember and know responses in recognition memory. *Psychonomic Bulletin & Review, 3,* 366–371.
Gardiner, J. M., & Parkin, A. J. (1990). Attention and recollective experience in recognition memory. *Memory & Cognition, 18,* 579–583.
Gregg, V. H., & Gardiner, J. M. (1991). Components of conscious awareness in a long-term modality effect. *British Journal of Psychology, 82,* 153–162.
Gregg, V. H., & Gardiner, J. H. (1994). Recognition memory and awareness: A large effect of study–test modalities on "know" responses following a highly perceptual orienting task. *European Journal of Cognitive Psychology, 6,* 131–147.
Guttentag, R. E., & Carroll, D. (1997). Recollection-based recognition: Word frequency effects. *Journal of Memory and Language, 37,* 502–516.
Hicks, J. L., & Marsh, R. L. (1999). Remember–know judgments can depend on how memory is tested. *Psychonomic Bulletin & Review, 6,* 117–122.
Hirshman, E., Fisher, J., Henthorn, T., Arndt, J., & Passannante, A. (2002). Midazolam amnesia and dual-process models of the word-frequency mirror effect. *Journal of Memory and Language, 47,* 499–516.
Hirshman, E., & Henzler, A. (1998). The role of decision processes in conscious recollection. *Psychological Science, 9,* 61–65.
Hirshman, E., & Lanning, K. (1999). Is there a special association between self judgements and conscious recollection? *Applied Cognitive Psychology, 13,* 29–42.
Hockley, W. E., & Consoli, A. (1999). Familiarity and recollection in item and associative recognition. *Memory & Cognition, 27,* 657–664.
Huron, C., Danion, J.-M., Giacomoni, F., Grangé, D., Robert, P., & Rizzo, L. (1995). Impairment of recognition memory with, but not without, conscious recollection in schizophrenia. *American Journal of Psychiatry, 152,* 1737–1742.
Inoue, C., & Bellezza, F. S. (1998). The detection model of recognition using know and remember judgments. *Memory & Cognition, 26,* 299–308.
Java, R. I., Gregg, V. H., & Gardiner, J. M. (1997). What do people actually remember (and know) in "remember/know" experiments? *European Journal of Cognitive Psychology, 9,* 187–197.
Java, R. I., Kaminska, Z., & Gardiner, J. M. (1995). Recognition memory and awareness for famous and obscure musical themes. *European Journal of Cognitive Psychology, 7,* 41–53.
Jones, T. C., & Roediger, H. L., III. (1995). The experiential basis of serial position effects. *European Journal of Cognitive Psychology, 7,* 65–80.
Joordens, S., & Hockley, W. E. (2000). Recollection and familiarity through the looking glass: When old does not mirror new. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1534–1555.
Khoe, W., Kroll, N. E. A., Yonelinas, A. P., Dobbins, I. A., & Knight, R. T. (2000). The contribution of recollection and familiarity to yes–no and forced-choice recognition tests in healthy subjects and amnesics. *Neuropsychologia, 38,* 1333–1341.
Kinoshita, S. (1995). The word frequency effect in recognition memory versus repetition priming. *Memory & Cognition, 23,* 569–580.
Knowlton, B. J., & Squire, L. R. (1995). Remembering and knowing: Two different expressions of declarative memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 699–710.
LeCompte, D. C. (1995). Recollective experience in the revelation effect: Separating the effects of recollection and familiarity. *Memory & Cognition, 23,* 324–334.
Lindsay, D. S., & Kelley, C. M. (1996). Creating illusions of familiarity in a cued recall remember/know paradigm. *Journal of Memory and Language, 35,* 197–211.
Mangels, J. A., Picton, T. W., & Craik, F. I. M. (2000). Attention and successful episodic encoding: An event-related potential study. *Cognitive Brain Research, 11,* 77–95.
Mäntylä, T., & Raudsepp, J. (1996). Recollective experience following suppression of focal attention. *European Journal of Cognitive Psychology, 8,* 195–203.
Mogg, K., Gardiner, J. M., Stavrou, A., & Golombok, S. (1992). Recollective experience and recognition memory for threat in clinical anxiety states. *Bulletin of the Psychonomic Society, 30,* 109–112.
Moscovitch, D. A., & McAndrews, M. P. (2002). Material-specific deficits in "remembering" in patients with unilateral temporal lobe epilepsy and excisions. *Neuropsychologia, 40,* 1335–1342.
Ochsner, K. N. (2000). Are affective events richly recollected or simply familiar? The experience and process of recognizing things past. *Journal of Experimental Psychology: General, 129,* 242–261.
Parkin, A. J., Gardiner, J. M., & Rosser, R. (1995). Functional aspects of recollective experience in face recognition. *Consciousness and Cognition, 4,* 387–398.
Parkin, A. J., & Russo, R. (1993). On the origin of functional differences in recollective experience. *Memory, 1,* 231–237.
Parkin, A. J., & Walter, B. M. (1992). Recollective experience, normal aging and frontal dysfunction. *Psychology and Aging, 7,* 290–298.
Payne, D. G., Elie, C. J., Blackwell, J. M., & Neuschatz, J. S. (1996). Memory illusions: Recalling, recognizing, and recollecting events that never occurred. *Journal of Memory and Language, 35,* 261–285.
Perfect, T. J., & Dasgupta, Z. R. R. (1997). What underlies the deficit in reported recollective experience in old age? *Memory & Cognition, 25,* 849–858.
Perfect, T. J., Williams, R. B., & Anderton-Brown, C. (1995). Age differences in reported recollective experience are due to encoding effects, not response bias. *Memory, 3,* 169–186.
Postma, A. (1999). The influence of decision criteria upon remembering and knowing in recognition memory. *Acta Psychologica, 103,* 65–76.
Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition, 21,* 89–102.
Rajaram, S. (1996). Perceptual effects on remembering: Recollective processes in picture recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 365–377.
Rajaram, S. (1998). The effects of conceptual salience and perceptual distinctiveness on conscious recollection. *Psychonomic Bulletin & Review, 5,* 71–78.
Rajaram, S., & Geraci, L. (2000). Conceptual fluency selectively influences knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1070–1074.
Reder, L. M., Nhouvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of the remember–know paradigm in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 294–320.
Schacter, D. L., Curran, T., Galluccio, L., Milberg, W. P., & Bates, J. F. (1996). False recognition and the right frontal lobe: A case study. *Neuropsychologia, 34,* 793–808.

(*Appendix continues*)

Schacter, D. L., Verfaellie, M., & Anes, M. D. (1997). Illusory memories in amnesic patients: Conceptual and perceptual false recognition. *Neuropsychology, 11,* 331–342.

Schacter, D. L., Verfaellie, M., & Pradere, D. (1996). The neuropsychology of memory illusions: False recall and recognition in amnesic patients. *Journal of Memory and Language, 35,* 319–334.

Smith, M. E. (1993). Neurophysiological manifestations of recollective experience during recognition memory judgments. *Journal of Cognitive Neuroscience, 5,* 1–13.

Strack, F., & Förster, J. (1995). Reporting recollective experiences: Direct access to memory systems? *Psychological Science, 6,* 352–358.

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

Verfaellie, M., Giovanello, K. S., & Keane, M. M. (2001). Recognition memory in amnesia: Effects of relaxing response criteria. *Cognitive, Affective, & Behavioral Neuroscience, 1,* 3–9.

Wagner, A. D., Gabrieli, J. D. E., & Verfaellie, M. (1997). Dissociations between familiarity processes in explicit recognition and implicit perceptual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 305–323.

Wippich, W. (1992). Implicit and explicit memory without awareness. *Psychological Research, 54,* 212–224.

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

Yonelinas, A. P., Dobbins, I., Szymanski, M. D., Dhaliwal, H. P., & 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., & 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.

Received May 19, 2003
Revision received October 17, 2003
Accepted October 22, 2003 ■