

Spotlighting the Probative Findings: Reply to Parks and Yonelinas (2007)

John T. Wixted
University of California, San Diego

J. T. Wixted (2007) argued that nearly every recent study that was designed to differentially test the unequal-variance signal-detection (UVSD) and dual-process signal-detection (DPSD) models has yielded findings that strongly support the UVSD model. C. M. Parks and A. P. Yonelinas (2007) did not address these studies in detail and instead focused mainly on issues that are tangential to the debate (e.g., measurement issues and neuroimaging). What few responses they did offer to the more relevant model-comparison studies did not effectively address the difficulties they posed for the DPSD model.

Keywords: signal-detection theory, dual-process theory, recollection, familiarity, ROC

In their reply to my article, Parks and Yonelinas (2007) addressed many issues, but they scarcely addressed the findings that speak directly to the relative validity of the unequal-variance signal-detection (UVSD) model versus the dual-process signal-detection (DPSD) model. These findings—which are the probative findings—are found primarily in studies that were specifically designed to pit the predictions of the two models against each other.

It is important to note at the outset that the competing accounts are both compatible with dual-process theory. The foundational disagreement between the DPSD model advanced by Yonelinas (1994) and the dual-process UVSD model advanced by Wixted (2007) concerns how the models conceptualize the recollection process. The DPSD model views recollection as a categorical phenomenon. According to this view, a test item either occasions conscious recollection of its prior occurrence on the list or it does not. When it does, it yields high confidence that the item is old. Parks and Yonelinas (2007) objected to the use of the phrase *all-or-none* to characterize that account of the recollection process, but it seems fair to characterize it as a *high-confident-or-none* account because, in the DPSD model, recollection is not associated with low or medium degrees of confidence. As Parks and Yonelinas (2007) said, on a 1-to-6 confidence scale, “under standard conditions recollection appears to be associated primarily with 6 responses” (p. 190).

And that is the crux of the issue. The dual-process UVSD model views recollection as a continuous process that is associated with low, medium, or high degrees of confidence, depending on the degree of recollection that is occasioned by the test item. The two models agree that familiarity is a continuous process, so the differences between them flow from this seemingly minor theoretical disagreement concerning the nature of recollection.

If the occurrence of recollection invariably yields high confidence that an item is old, then it makes sense to assume, as the DPSD model does (and as all prior dual-process models have), that individual recognition decisions are based either on recollection or

on familiarity. That is, according to this view, participants would quite reasonably base their decisions solely on recollection whenever possible because that process alone would be sufficient to support high confidence that the item is old. And they would base their decision solely on familiarity whenever recollection failed because, under those conditions, that would be the only option available to them.

But the logic changes if one allows for the possibility that recollection, like familiarity, is a continuous process that is associated with varying degrees of confidence and varying degrees of accuracy. In that case, responding based either on recollection or on familiarity would not make a great deal of sense. Unless they were perfectly correlated, the two processes combined would make for a more compelling memory strength variable than either one alone. In the combined model, one’s degree of certainty that an item appeared on a list is much like a juror’s degree of certainty that a defendant is guilty. In both cases, multiple sources of evidence are combined into an aggregate variable upon which the decision is based. According to this way of thinking, dual-process theory and signal-detection theory—the twin peaks of recognition memory for decades—are inherently reconcilable. Moreover, as I describe next, the most relevant data strongly support this way of thinking over the DPSD model.

Describing the Receiver-Operating Characteristic (ROC)

Any viable quantitative account of recognition memory must be able to characterize the ROC at least as well as the UVSD model. In their reply to my article, Parks and Yonelinas (2007) relegated their discussion about this critical issue to a footnote (Footnote 4). In that footnote, they suggested that the two models offer comparable fits of the ROC, but a more complete summary of the relevant evidence shows otherwise. It is certainly true that the DPSD model fits ROC data reasonably well, but that is because it incorporates a signal-detection component. In fact, the UVSD and DPSD models become identical as the slope of the z -ROC approaches 1.0. The question of interest is how the DPSD model’s departure from signal-detection theory (i.e., its threshold recollection component) affects its ability to fit the data. The answer is that to the extent that it deviates from the standard detection model, it deviates from the ROC data as well.

Correspondence concerning this article should be addressed to John T. Wixted, Department of Psychology, 0109, University of California, San Diego, La Jolla, CA 92093. E-mail: jwixted@ucsd.edu

Because they are partially overlapping models that both fit ROC data well, it is not surprising that some studies lack the power to distinguish between them. However, studies that do have the power to detect a significant difference tell a compelling story. As noted by Parks and Yonelinas (2007), two studies have reported an advantage for the DPSD model over the UVSD model. First, Yonelinas, Dobbins, Szymanski, Dhaliwal, and King (1996) reported that in one of four conditions, the DPSD model offered a significantly better fit than the UVSD model (whereas the other three conditions yielded null results). Second, Howard, Bessette-Symons, Zhang, and Hoyer (2006), using complex travel scenes as stimuli, found a significant advantage for the DPSD model in both of their conditions (young and old participants). Thus, in total, three experimental conditions in which a significant difference was detected favored the DPSD model over the UVSD model.

Against these three conditions are many more conditions that show the opposite result. Heathcote (2003), in what is by far the most comprehensive direct comparison between the two models, found a significant advantage for the UVSD model in 23 different conditions across four experiments. The DPSD did not offer a better fit in a single condition of that large study.¹ In addition, Rotello, Macmillan, Hicks, and Hautus (in press) found that the UVSD model offered a better fit in both of the conditions they reported, Dougal and Rotello (in press) found that the detection model fit better in the two conditions of their experiment, and Healy, Light, and Chung (2005) found that the detection model offered a better fit in the two conditions they examined (young and old). Thus, of the studies that have detected a difference, an overwhelming advantage is evident for the UVSD model (29 conditions against 3).

All of these studies involved an examination of individual participant ROCs, but group ROC analyses tell the same story. Heathcote (2003) examined group ROC data previously reported by Glanzer, Kim, Hilford, and Adams (1999). Yonelinas (1999b) suggested that the two models fit those group data equally well, but that argument was based on the fact that both models accounted for an extremely high percentage of the data variance. Heathcote (2003) found that although both models fit the Glanzer et al. (1999) data well, the detection model fit best in 10 out of 10 conditions (an outcome unlikely to be due to chance). Similarly, Wais, Wixted, Hopkins, and Squire (2006) recently found that the detection model fit their group ROC data best in 7 out of 7 conditions. Moreover, as Heathcote (2003) reported, the disadvantage of the DPSD model was greater as the slope of the ROC decreased. The DPSD model accounts for a decreased ROC slope by assuming a greater role for the recollection parameter. Thus, the DPSD model's deficiency became more evident to the extent that the recollection process was involved in the fit. This fact underscores the main theme of my reply: The way in which the DPSD model conceptualizes the recollection process is inaccurate.

Predicting Forced-Choice Performance From Old–New Performance

Smith and Duncan (2004) reported that the UVSD model does a very good job of predicting a participant's forced-choice recognition performance on the basis of a fit of the UVSD model to the participant's old–new ROC data. The DPSD model, by contrast,

does a much poorer job, and the main failing is clearly the recollection parameter (which offers no predictive validity at all).

In response to this key finding, Parks and Yonelinas (2007), in their Footnote 6, suggested that this study was problematic because participants were instructed to spread their responses evenly across the confidence categories. As they put it,

What this means is that participants could only place a maximum of about 33% of the old items (47/140) in the high-confidence bin. If a participant recollected more than 33% of the old items, which is likely given the high level of observed performance, then they would have been forced to distribute the recollected items across lower confidence bins. (p. 198)

But Smith and Duncan (2004) noted that they did not enforce compliance with this instruction, and it is quite clear from their results that their participants did not feel bound by it. If they were bound by it, the maximum possible recollection estimate from the DPSD model would be .33, but the actual recollection estimates were evenly distributed from 0 to .60 (as seen in their Figure 4C), with approximately half the participants falling above .30 and half falling below.

Even if participants had been constrained, the results would still not make sense from the point of view of the DPSD model and would still validate the UVSD model. Parks and Yonelinas (2007) argued that had participants been constrained, it would "lead the DPSD model to mis-fit the data" and would "artificially make it look like the UVSD model is fitting better than it really does" (p. 198). But this experiment was not primarily about fitting the data. The UVSD model did fit the data better than the DPSD model (as usual), but the key finding had to do with each model's *predictive* validity. It is hard to imagine how, if the DPSD model is fundamentally correct, asking participants to spread their responses across confidence categories would not only cause the DPSD model to fail but would also cause the UVSD model to succeed in such spectacular fashion when it comes to predicting forced-choice performance. The UVSD model showed impressive predictive validity (accounting for 66% of the variance across individuals), a finding that would appear to suggest that the model's underlying assumptions are accurate.

¹ Parks and Yonelinas (2007) discounted the findings reported by Heathcote (2003) on the basis of an apparent misreading of what he reported. They claimed that in most cases Heathcote (2003) found that the recognition ROCs exhibited an inverted U shape in z space, which is problematic for both models. Actually, he found that anomaly only in a small subset of cases for which the UVSD model did not provide a good fit. In Heathcote's Experiment 3, for example, the UVSD model was judged to mis-fit the ROC data in only 19 of 238 cases. Of those 19, 12 exhibited an inverted U shape in z space, which is the anomaly that Parks and Yonelinas (2007) suggested was observed most of the time. In fact, it was rarely observed, and his findings certainly cannot be discounted on that basis. Moreover, that the z -ROC might occasionally exhibit this anomaly is easily understood and does not require any new theoretical considerations. In particular, the inverted U shape in z space would occur if participants occasionally make random guesses, which, occasionally, they surely do. As Heathcote said about these cases: "Hence, the observed minor deviations from a linear z -ROC function are consistent with a small percentage of guessing responses for some participants" (pp. 1216–1217).

Recollection and the Remember–Know Procedure

Many assume that remember–know judgments map onto recollection and familiarity, and this seems especially true of those working on the neuroanatomical basis of recognition memory (e.g., Aggleton & Brown, 2006). However, much evidence supports Donaldson's (1996) signal-detection interpretation of remember–know judgments instead (Dunn, 2004; Wixted & Stretch, 2004). Donaldson (1996) argued that remember responses reflect strong memories and know responses reflect weaker memories. According to the dual-process UVSD model proposed by Wixted (2007), recollection and familiarity both play a role in both strong and weak memories. If so, then high-strength remember responses correspond to high degrees of recollection and familiarity (on average) and low-strength know responses correspond to low degrees of recollection and familiarity (on average).

Parks and Yonelinas (2007) stated that “[remember–know] reports can be used to assess recollection and familiarity, but they do not provide process-pure measures of these processes” (p. 193). They addressed the measurement issue in considerable detail, but my argument is concerned with the widely held idea that remember–know judgments are based on recollection and familiarity, respectively (not that they provide pure measures of those processes). In my critique, I did not consider the measurement issue at all. If the core idea that individual remember and know judgments correspond to recollection and familiarity is wrong, then the debate about whether a measure of familiarity is provided by simply counting up know responses or by correcting the number of know responses on the basis of the number of remember responses is beside the point. If the foundational assumption is incorrect, then it does not matter what exact formula is used to derive a putative estimate of recollection or familiarity.

Parks and Yonelinas (2007) said,

Given the compelling subjective experiences of recollection and familiarity that we have all had, it is hard to believe that people are unable to report when they retrieve qualitative information and when they recognize something as old but cannot recall anything about it. (p. 194)

My claim is not that people are constitutionally unable to do this. Instead, my claim is that these process-pure extremes are rare exceptions, not the norm. Normally, both processes contribute to the decision, so asking participants to indicate which process was involved is like asking a juror which one of two pieces of evidence, fingerprint evidence or fiber evidence, was involved in their decision to find the defendant guilty. The question can be asked, but the assumption that underlies the question may be wrong. Because the remember–know instructions do not correspond to typical experience, participants theoretically solve the problem by simply setting different strength criteria.

Parks and Yonelinas (2007) spent a good deal of time discussing the instructions that might be needed so that participants will be sure to say “remember” only when recollective details are retrieved (e.g., by making it clear that they may be asked about those details). But the probative empirical consideration—one that Parks and Yonelinas chose not to address—is that the evidence clearly suggests that know responses are not familiarity based. The key finding is that source recollection accuracy is almost invariably above chance for know responses (which are ostensibly based on

familiarity). Recollection accuracy is even higher for remember responses, which makes sense according to both the DPSD model and the UVSD model, but it should be at chance for familiarity-based know responses according to the DPSD model. Instead, it is clearly above chance. This suggests that know responses reflect weaker memories (i.e., they involve less recollection) than remember responses, which is how it must be according to the dual-process UVSD model.

Recollection and Confidence

The same problem arises with respect to confidence judgments, but again, the issue is simply not addressed by Parks and Yonelinas (2007) in their reply to my article. The relationship between confidence and recollection lies at the very heart of the disagreement. The dual-process UVSD model requires that recollection accuracy be related in graded fashion to increases in confidence in the old decision. That is, as confidence in the old decision increases, accuracy in a subsequent source recollection question should increase accordingly. The DPSD model, by contrast, requires that recollection be associated with high-confidence old decisions only. Every single application of the DPSD model has assumed that recollection-based responses result in high confidence (e.g., 6 on a 6-point confidence scale). As such, high-confidence old decisions should be associated with accurate source recollection, but less-than-high-confidence old decisions should be associated with chance performance on the recollection test. This very pattern was reported by Yonelinas et al. (1996), but the evidence from multiple studies conducted since that time (the results of which are summarized in Table 3 of Wixted, 2007) show quite clearly that partial recollection is associated with less-than-high-confidence old decisions. Specifically, when participants express low confidence in an old decision, subsequent source accuracy is slightly above chance. When they express medium confidence, subsequent source accuracy is higher than that, and when they express high confidence, source accuracy is higher still. In other words, confidence in an old–new decision and source recollection accuracy are related in graded fashion (as the dual-process UVSD model requires), not in categorical fashion (as the DPSD model requires). This is a critically important, probative, result that does not lend itself to a ready explanation if the DPSD model is correct.

Relational Recognition Tests

As Parks and Yonelinas (2007) noted, source memory tests and associative recognition tests are of special interest in this debate because they are widely assumed to preferentially tap the recollection process (which is the process that lies at the heart of the disagreement). For the sake of brevity, I focus mainly on source memory in this section. In a source memory procedure, participants are first asked whether the item appeared on the list (old or new) and are then asked to recollect the item's source (male voice or female voice). Item familiarity contributes to the old–new decision (i.e., old items are more familiar than new items, on average) but is of no help on the source question because the items from both sources are, by design, equally familiar, on average. As such, recollection is needed to identify the source, and if recollection is a threshold process, the ROC should be linear. However, the

overwhelming weight of evidence shows that relational ROCs are typically curvilinear, though not as curvilinear as they should be according to the UVSD model.

Does the DPSD Model Predict a Linear Source Memory ROC?

In their reply to my article, Parks and Yonelinas (2007) took issue with my claim that the DPSD model predicts linear source memory and associative recognition ROCs. In particular, they said,

Nonetheless, Wixted (2007) took the findings of curvilinear ROCs in relational tests as evidence that the DPSD model fails. However, this is exactly what the model predicts if one assumes that recollection and familiarity contribute to performance. Curvilinear ROCs in relational recognition are only inconsistent with the DPSD model if one makes the incorrect assumption that these tests provide a process-pure measure of recollection. (p. 195)

It is certainly true that the DPSD model can predict curvilinear ROCs under certain circumstances (e.g., if the experiment is designed to make items from one source more familiar than the items from the other source), but the idea that the model predicts linear ROCs under typical circumstances is not mine. For example, in his article on associative recognition ROCs, Yonelinas (1997) stated that

The results of Experiments 1–3 show that ROCs for item judgments are curvilinear, but that ROCs for associative judgments are linear. These results were observed with the average and individual subject ROCs. The curvilinear item ROCs were similar to those found in many previous studies of item recognition and are to be expected if recognition judgments rely on a combination of recollection and familiarity. *In contrast, the linear associative ROCs that were observed were predicted on the basis of the assumption that associative recognition judgments rely on recollection.* (p. 759, italics added)

In addition, in his article on source memory ROCs, Yonelinas (1999a) stated that

Recognition ROCs were curvilinear in probability space, and they were generally linear in z -space. *In contrast, source ROCs were generally linear in probability space and U-shaped in z -space. These results were in agreement with the predictions of the dual-process model, and the observed ROCs were fit well by the model's equations.* In the recognition tests, in which participants were expected to rely on both recollection and familiarity, the ROCs were curved and asymmetrical. In contrast, in the source tests, in which participants were expected to rely primarily on recollection (i.e., Experiments 1–3), the ROCs were linear. (p. 1428, italics added)

Initial findings suggesting that associative recognition and source memory ROCs really are linear—as uniquely predicted by the DPSD model—legitimized the model more than any other finding to date. Thus, the fact that a large number of subsequent studies have found these ROCs to be typically curvilinear is directly relevant to this debate.

The existence of curvilinear relational ROCs forces the DPSD model to assume that familiarity plays a role in discriminating the item's source even when it is reasonable to assume that the two sets of items are equally familiar. If they are not equally familiar for some reason (e.g., if words spoken in a female voice are more familiar than words spoken in a male voice), then the source

discrimination can be aided by relying on the familiarity process on those occasions when recollection fails. Doing so would, according to the DPSD model, impart some degree of curvilinearity to the source ROC. However, because source memory studies are usually specifically designed in such a way that the items from the two sources are equally familiar, the appeal to differences in source familiarity in order to account for curvilinear source ROCs is not intuitively compelling. More to the point, the claim that the items from one source are more familiar than the items from the other source is testable because it makes a prediction about how easily the items from each source can be discriminated from new items. That is, the ostensibly more familiar Source A items should be easier to discriminate from new items than the ostensibly less familiar Source B items. That test appears among the quantitative analyses of the conditional source ROC that I present next.

Conditional Source ROCs

In my article, I claimed that an analysis performed by Slotnick and Dodson (2005) on the Yonelinas (1999a) source ROC data decisively settled the issue in favor of the UVSD model. Parks and Yonelinas (2007) disagreed, but a more detailed look at the issue strongly reinforces my original claim. An illuminating way to analyze the source ROC is to partition the data as a function of confidence in the old–new decision that precedes the source decision, as Slotnick and Dodson (2005) did. This contrasts with the method Yonelinas (1999a) used, which was to construct a single source ROC using all of the data regardless of the level of confidence expressed in the initial old–new decision.

According to the DPSD model, if source recollection occurs during the old–new stage, a high-confidence old decision will be made (i.e., recollection yields high confidence). The DPSD model allows for the possibility that some additional high-confidence old–new decisions can be based on high item familiarity, but all old–new decisions made with lower confidence are assumed to be based solely on familiarity. Imagine 100 old decisions, 60 of which were made with high confidence. Of those 60 high-confidence old decisions, imagine that 40 were based on source recollection and 20 on familiarity. If a source ROC is constructed using all 100 items, then, if the DPSD model is accurate, it will return an estimate of .40 for recollection because 40 out of 100 decisions were recollection based. It would return this estimate if the model were fit either to the curvilinear old–new ROC or to the more linear source ROC. In practice, the old–new and source recollection estimates are similar but not identical (with the old–new estimate being slightly higher), perhaps because some additional nonsource recollection occurs during the old–new stage (Yonelinas, 1999a). At this level of analysis, the results largely accord with the DPSD model.

If we now construct a conditional source ROC using only the items that received a high-confidence old decision, the DPSD model must return a higher recollection estimate because now 40 of the 60 responses making up the ROC are recollection based. Ideally, the recollection estimate derived from this fit would be .67. Because a higher proportion of the responses in the conditional ROC are theoretically based on threshold recollection, the UVSD model—which includes no threshold processes—should have an even harder time fitting the data. The DPSD model, by contrast, should continue to fit very well. Slotnick and Dodson

(2005) showed that the more purely recollection-based conditional ROC was unexpectedly quite curvilinear and, unlike the collapsed ROC, was very well described by the UVSD model. Figure 1 shows the collapsed ROC that is made up of all responses as well as the conditional ROC that is made up of only source decisions that followed a high-confidence old decision. The curvilinearity of the conditional ROC is readily apparent. This outcome makes no sense if the DPSD model is correct, and it suggests that the inability of the UVSD to accurately describe the collapsed source ROC may not mean what it initially appeared to mean.

The collapsed and conditional ROCs in Figure 1 were constructed using data originally reported by Yonelinas (1999a) and reproduced in more detail by Slotnick and Dodson (2005). Parks and Yonelinas (2007) took me to task for not actually fitting the DPSD model to the conditional source ROC data. In response, I fit both the DPSD model and the UVSD model to the collapsed and conditional source memory ROC data shown in Figure 1 (using maximum likelihood estimation), and the results are illuminating.²

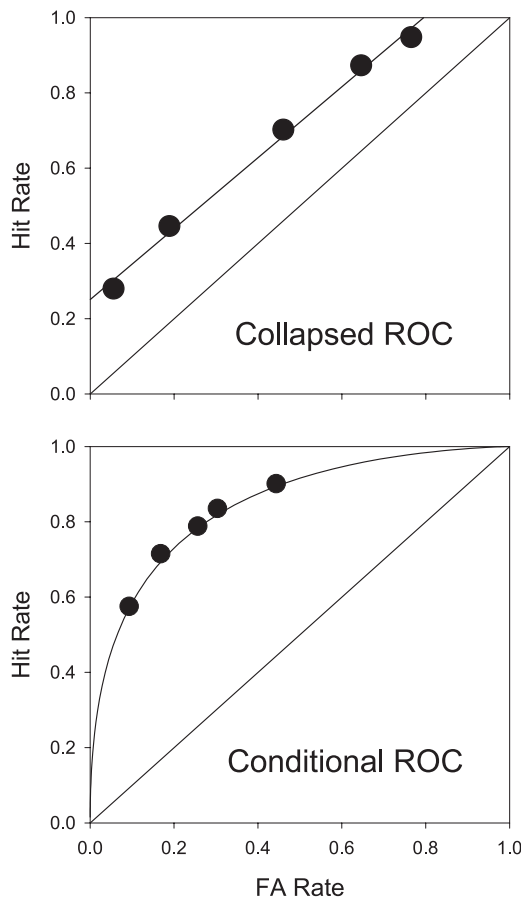


Figure 1. The top panel shows the collapsed source memory receiver-operating characteristic (ROC) data (along with the best-fitting straight line), and the bottom panel shows the conditional source ROC using only source decisions that were preceded by a high-confidence old decision (along with the best-fitting fit of the unequal-variance signal-detection model). The confidence ratings are from Yonelinas (1999a, Experiment 2) and were taken from Table 1 of Slotnick and Dodson (2005). FA = false alarm.

The full DPSD model has three significant parameters: R_m (probability of correctly recollecting that the item was presented in a male voice), R_f (probability of correctly recollecting that the item was presented in a female voice), and d' (the equal-variance familiarity parameter). If the source ROC (either the collapsed or the conditional) is linear, the d' parameter will equal 0 (indicating no role for familiarity, according to the DPSD model). If it exhibits curvilinearity, the d' parameter will be greater than zero (indicating that familiarity does play a role).

The DPSD model fit the collapsed source ROC well, $\chi^2(2, 3840) = 4.16$, with $R_m = .14$, $R_f = .19$, and $d' = 0.27$. Although the ROC is almost linear, these values suggest that the ROC was slightly curvilinear (with $d' = 0.27$), which, in turn, suggests a slight role for familiarity. The two-parameter version of the DPSD model (with R_m constrained to equal R_f), which is the version that Yonelinas (1999a) used, also fit reasonably well, $\chi^2(3, 3840) = 9.83$. The two-parameter UVSD model, by contrast, does a very poor job of fitting this nearly linear ROC, $\chi^2(3, 3840) = 62.9$. This is essentially what Yonelinas (1999a) reported, and on the basis of that analysis, it was once reasonable to assume that these data pose real problems for the UVSD model and validate the DPSD model (despite the slight curvilinearity that the DPSD model would not necessarily expect). The almost linear source ROC is U shaped in z space, and I agree with Parks and Yonelinas (2007) that this result was predicted a priori by the DPSD model (a point they drive home in their Figure 2).

I next fit the conditional source memory ROC, which was based on high-confidence old decisions only. As indicated above, these are the decisions that theoretically contain all of the recollection-based source decisions. In fact, theoretically, these are the only responses that contribute meaningfully to the collapsed source ROC function (because all other responses contain no recollection and should, when considered separately, fall along the diagonal). Thus, the model predicts that the recollection parameters of the DPSD model will increase substantially and that the ROC should remain as linear as the collapsed ROC. However, as I show next, precisely the opposite happens when the conditional ROC is analyzed.

The three-parameter DPSD model fit the conditional source ROC reasonably well, $\chi^2(2, 1660) = 4.96$, but the parameters tell a theoretically nonsensical story. Far from increasing, the recollection parameters decreased substantially, with $R_m = 0$ (down from .14) and $R_f = .14$ (down from .19). Moreover, the familiarity parameter, d' , exhibited a fivefold increase (from 0.27 to 1.32). These seemingly odd parameter changes occurred because the conditional ROC—which every model agrees is largely recollection based—is clearly curvilinear. The only way the DPSD model can interpret curvilinearity is to assume that familiarity played a large role. But the only way that familiarity could play a large role in the source discrimination is if there was a large difference in familiarity, on average, for items presented in a male voice versus items presented in a female voice (e.g., female voices very familiar, male voices very unfamiliar). During the old–new stage, that difference would give rise to a difference in discriminability as a function of source, with the more familiar items being recognized better than the less familiar ones. However, a more detailed look at

² I thank Caren Rotello for suggesting these analyses.

the male–female data shows no apparent difference in old–new discriminability. More specifically, fitting the DPSD model separately to the male and female old–new ROCs yields familiarity d' estimates of 0.60 and 0.62, respectively, and threshold recollection estimates of .32 and .34, respectively. In other words, according to the DPSD model itself, familiarity and recollection were equated for the male and female items. Under such conditions, only recollection can aid the source memory discrimination. These estimates are consistent with what is usually assumed to be true on the basis of the design of source memory tests: Familiarity offers no assistance when it comes to identifying the source. Even so, the conditional source ROC is quite curvilinear.

An important additional test is provided by fitting a four-parameter model to the conditional source ROC data. In this four-parameter model, a new parameter, s , is added to the DPSD model to allow for the possibility of a UVSD component (whereas the model ordinarily assumes an equal-variance component). The starting value of this parameter is 1.0, and any deviation from 1.0 implies unequal variance for the signal-detection component. If the DPSD model is valid, this parameter will remain close to 1.0 because that model assumes that familiarity is an equal-variance process.

When the four-parameter model is fit to the collapsed source memory ROC, the unequal-variance parameter remains close to 1.0, and its presence does not significantly (or even minimally) improve the fit: $\chi^2(1, 3840) = 4.05$, with $R_m = .20$, $R_f = .13$, $d' = 0.27$, and $s = 1.035$. This is virtually the same result one obtains from fitting the three-parameter model, and it is just what should happen according to the DPSD model. However, when the same approach is used for the recollection-based conditional source ROC, a very different result obtains. The recollection parameters now drop away (i.e., both equal 0), leaving only the two parameters of the signal-detection component. The fit is excellent, $\chi^2(3, 1660) = 4.37$, with $d' = 1.52$, and $s = 0.91$. In other words (and this is the key point), the data are fully explained by the two-parameter UVSD model. The recollection parameters add nothing to the fit of this recollection-based ROC.

Note that this fit is as fair as it could be to the DPSD model. It uses data reported by Yonelinas (1999a) to support that model. It uses only data that involved a high-confidence old decision, which all models (especially the DPSD model) agree is where source recollection is concentrated. It does not insist that the data be linear, as originally predicted by the DPSD model, because it allows for a familiarity component in source decisions (odd as that might seem). But when the four-parameter fit is performed, a two-parameter result is returned: R_m and R_f both drop to 0 (i.e., they drop out), leaving only the two parameters of the UVSD model to fit the data.

If the DPSD model is correct, the ability of the UVSD model to fit the conditional ROC data is problematic because one needs to implausibly assume that familiarity played an especially large role (and recollection played a reduced role) when all relevant models suggest that recollection played an especially large role instead. A simpler interpretation is provided by the dual-process UVSD model: The conditional ROC does consist of recollection-based data, but recollection is a continuous process, not a threshold process (hence, the curvilinearity).

Why, then, does the DPSD model have such a clear advantage when fitting the nearly linear collapsed source ROC shown in the

upper panel of Figure 1? Because, as Slotnick and Dodson (2005) demonstrated, the collapsed ROC includes many responses that, when examined separately, fall along the linear diagonal of the ROC. That is, these are items for which no source information is available. The simplest explanation for why this might happen is based on the fact that during study, item information must be encoded before source information can be encoded. For some items, no source information will be encoded if item processing did not proceed far enough to allow that to happen. These items will fall along the diagonal in the source ROC, thereby imparting apparent linearity to the collapsed source ROC.

Conclusion

The key point is that studies that directly pit the predictions of the DPSD model and the UVSD model against each other (which are the studies that provide the probative data) are nearly unanimous in their endorsement of the UVSD model. In their response to my article, Parks and Yonelinas (2007) paid scant attention to these studies and instead asserted that many studies in the neuroscience literature—studies that were not designed to differentiate the two models—strongly support the DPSD model. In this regard, they said,

Studies using fMRI have provided additional support for the DPSD model. Although Wixted (2007) cited a few neuroimaging studies in his review, he failed to point this out. For example, fMRI studies using [remember–know], relational recognition, and ROC confidence methods have indicated that recollection is consistently associated with hippocampal activation. . . . (p. 196)

These neuroimaging studies do not offer probative findings with regard to the question of whether the DPSD model or the UVSD model is more valid because they were not designed to do so. Instead, imaging results are typically interpreted through the lens of a particular psychological model. Similar considerations apply to this statement:

Studies using [remember–know], relational, and ROC methods have indicated that patients with damage including the hippocampus and the surrounding medial temporal cortex have deficits in recollection as well as smaller, but consistent, deficits in familiarity. . . . (p. 196)

Again, with a few exceptions (e.g., Wais et al., 2006), most of the studies they cited presuppose the validity of the DPSD model, which is often used to derive quantitative estimates of recollection and familiarity.

The search for the neuroanatomical correlates of recognition memory must be guided by some psychological model, and the model that usually serves that role is the DPSD model (or a close relative of it) despite that it has been largely invalidated in the psychological literature. The UVSD model offers a more viable model, and as described by Wixted (2007), this model suggests that the procedures that have been devised to separate recollection from familiarity have succeeded only in separating strong memories (high in both recollection and familiarity) from weak memories (lower in both recollection and familiarity). Alternative methods to disentangle the two processes are suggested by the dual-process UVSD model, but it turns out to be a bit more complicated than simply asking for remember or know judgments (Wixted, 2007). Complicated or not, the search for the neuroanatomical

correlates of recognition memory should be guided by a psychologically valid model.

References

Aggleton, J. P., & Brown, M. W. (2006). Interleaving brain systems for episodic and recognition memory. Trends in Cognitive Sciences, 10, 455-463.
Donaldson, W. (1996). The role of decision processes in remembering and knowing. Memory & Cognition, 24, 523-533.
Dougal, S., & Rotello, C. M. (in press). "Remembering" emotional words is based on response bias, not recollection. Psychonomic Bulletin & Review.
Dunn, J. C. (2004). Remember-Know: A matter of confidence. Psychological Review, 111, 524-542.
Glanzer, M., Kim, K., Hilford, A., & Adams, J. K. (1999). Slope of the receiver-operating characteristic in recognition memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 25, 500-513.
Healy, M. R., Light, L. L., & Chung, C. (2005). Dual-process models of associative recognition in young and older adults: Evidence from receiver operating characteristics. Journal of Experimental Psychology: Learning, Memory, and Cognition, 31, 768-788.
Heathcote, A. (2003). Item recognition memory and the ROC. Journal of Experimental Psychology: Learning, Memory, and Cognition, 29, 1210-1230.
Howard, M. W., Bessette-Symons, B., Zhang, Y., & Hoyer, W. J. (2006). Aging selectively impairs recollection in recognition memory for pictures: Evidence from modeling and receiver operating characteristic curves. Psychology and Aging, 21, 96-106.
Parks, C. M., & Yonelinas, A. P. (2007). Moving beyond pure signal-detection models: Comment on Wixted (2007). Psychological Review, 114, 188-202.
Rotello, C. M., Macmillan, N. A., Hicks, J. L., & Hautus, M. (in press). Interpreting the effects of response bias on remember-know judgments using signal-detection and threshold models. Memory & Cognition.

Slotnick, S. D., & Dodson, C. S. (2005). Support for a continuous (single-process) model of recognition memory and source memory. Memory & Cognition, 33, 151-170.
Smith, D. G., & Duncan, M. J. J. (2004). Testing theories of recognition memory by predicting performance across paradigms. Journal of Experimental Psychology: Learning, Memory, and Cognition, 30, 615-625.
Wais, P. E., Wixted, J. T., Hopkins, R. O., & Squire, L. R. (2006). The hippocampus supports both the recollection and the familiarity components of recognition memory. Neuron, 49, 459-466.
Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. Psychological Review, 114, 152-176.
Wixted, J. T., & Stretch, V. (2004). In defense of the signal detection interpretation of remember/know judgments. Psychonomic Bulletin & Review, 11, 616-641.
Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. Journal of Experimental Psychology: Learning, Memory, and Cognition, 20, 1341-1354.
Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. Memory & Cognition, 25, 747-763.
Yonelinas, A. P. (1999a). The contribution of recollection and familiarity to recognition and source memory: An analysis of receiver operating characteristics and a formal model. Journal of Experimental Psychology: Learning, Memory, and Cognition, 25, 1415-1434.
Yonelinas, A. P. (1999b). Recognition memory ROCs and the dual-process signal detection model: Comment on Glanzer, Kim, Hilford, and Adams (1999). Journal of Experimental Psychology: Learning, Memory, and Cognition, 25, 514-521.
Yonelinas, A. P., Dobbins, I., Szymanski, M. D., Dhaliwal, H. S., & King, L. (1996). Signal-detection, threshold, and dual-process models of recognition memory: ROCs and conscious recollection. Consciousness and Cognition, 5, 418-441.

Received September 21, 2006
Revision received September 28, 2006
Accepted September 29, 2006

Statement of Ownership, Management, and Circulation
1. Publication Title: Psychological Review
2. Publication Number: 10.1037/0033-2909.114.3.188
3. Filing Date: October 2006
4. Issue Frequency: Quarterly
5. Number of Issues Published Annually: 4
6. Annual Subscription Price: \$67 indiv \$142 inst
7. Complete Mailing Address of Known Office of Publication: 750 First Street, N.E., Washington, D.C. 20002-4242
8. Complete Mailing Address of Headquarters or General Business Office of Publisher: 750 First Street, N.E., Washington, D.C. 20002-4242
9. Full Names and Complete Mailing Addresses of Publisher, Editor, and Managing Editor: American Psychological Association, 750 First Street, N.E., Washington, D.C. 20002-4242

17. Extent and Nature of Circulation
a. Total Number of Copies (Net press run): 4800
b. Paid and/or Requested Circulation: 2492
c. Total Paid and/or Requested Circulation: 3516
d. Total Distribution (Sum of 15c and 15d): 3730
e. Total (Sum of 15e and 15f): 4800
18. Publication of Statement of Ownership: Yes
19. Signature and Title of Editor, Publisher, Business Manager, or Owner: Barbara Spruffl, Director
20. Tax Status: None
21. Has Not Changed During Preceding 12 Months: Yes
22. Has Changed During Preceding 12 Months: No