

Variation in the standard deviation of the lure rating distribution: Implications for estimates of recollection probability

Stephen Dopkins¹ · Kaitlin Varner¹ · Darin Hoyer¹

© Psychonomic Society, Inc. 2017

Abstract In word recognition semantic priming of test words increased the false-alarm rate and the mean of confidence ratings to lures. Such priming also increased the standard deviation of confidence ratings to lures and the slope of the z-ROC function, suggesting that the priming increased the standard deviation of the lure evidence distribution. The Unequal Variance Signal Detection (UVSD) model interpreted the priming as increasing the standard deviation of the lure evidence distribution. Without additional parameters the Dual Process Signal Detection (DPSD) model could only accommodate the results by fitting the data for related and unrelated primes separately, interpreting the priming, implausibly, as decreasing the probability of target recollection (DPSD). With an additional parameter, for the probability of false (lure) recollection the model could fit the data for related and unrelated primes together, interpreting the priming as increasing the probability of false recollection. These results suggest that DPSD estimates of target recollection probability will decrease with increases in the lure confidence/evidence standard deviation unless a parameter is included for false recollection. Unfortunately the size of a given lure confidence/evidence standard deviation relative to other possible lure confidence/evidence standard deviations is often unspecified by context. Hence the model often has no way of estimating false recollection probability and thereby correcting its estimates of target recollection probability.

Keywords Recollection · Familiarity · ROC curve · Unequal variance signal detection · Dual process signal detection

Most writers agree that recognition can occur through two different processes: a stimulus can be recognized because it is familiar or because it is recollected as having occurred previously in a certain context (Wixted, 2007; Yonelinas & Parks, 2007). What is unclear is the magnitude of recollection's role in recognition (Ratcliff & Starns, 2009). Item recognition – perhaps the simplest form of recognition – is important for addressing this question. What role does recollection play here?

An influential proposal seeks to answer this question by examining the receiver operating characteristic (ROC) function. This function plots the proportion of *hits* (correctly-detected *targets*) against the proportion of *false alarms* (incorrectly-detected *lures*) for multiple criteria. Critically for the present discussion, the ROC function for item recognition is typically asymmetrical. This is best seen when the proportions of hits and false alarms are transformed to z scores; the z-ROC function is typically linear; the asymmetry is seen in that the slope of the function is less than 1 (Ratcliff, Sheu, & Gronlund, 1992).

The Dual Process Signal Detection (DPSD) model infers the level of target recollection in item recognition from the degree of asymmetry in the ROC function. The model assumes that item recognition reflects familiarity assessment and recollection. Familiarity assessment occurs through a signal detection process. Targets and lures are associated with distinct distributions of mnemonic evidence. The observer emits a positive familiarity-based response to a test item when its mnemonic evidence exceeds a criterion. Recollection occurs through a threshold process. If the mnemonic evidence for a test item exceeds a particular high value, the item is

✉ Stephen Dopkins
dopkins@gwu.edu

¹ Psychology Department, George Washington University, 2125 G Street NW, Washington, DC 20052, USA

recollected. The target and lure evidence distributions for the familiarity process have equivalent standard deviations. However, the target evidence distribution for the overall recognition process has a larger standard deviation than the lure evidence distribution for that process because recollection plays a greater role in judgments to targets than to lures. The slope of the z-ROC function reflects the ratio of the target and lure evidence standard deviations for the recognition process. Hence the level of target recollection can be inferred from the slope of the z-ROC function (Koen & Yonelinas, 2010).

The DPSD interpretation of ROC data has been influential, in neuropsychology and neuroscience as well as in human memory (Lipton & Eichenbaum, 2012; Schoemaker, Gauthier, & Pruessner, 2014). It has not gone unchallenged, however. According to the Unequal Variance Signal Detection (UVSD) model, item recognition occurs through a unitary signal detection process (which may reflect familiarity and recollection information). The target distribution has a larger standard deviation than the lure distribution and the slope of the z-ROC function reflects the ratio of the two standard deviations. However, this ratio is not tied to the level of target recollection (Wixted, 2007).

Past work has sought to contrast the DPSD and UVSD models by manipulating the target evidence standard deviation (Jang, Mickes, & Wixted, 2012; Koen and Yonelinas, 2010; Starns, Rotello, & Ratcliff, 2012). The present study sought to contrast the models by manipulating the lure evidence standard deviation. During the study phase of each block of the experiment, the participant read a list of words. During the test phase of each block the participant responded to a series of word pairs, making (1) a recognition judgment to the *prime word*, (2) a recognition judgment to the *test word*, and (3) a confidence rating regarding the recognition judgment to the test word. The prime word and test word were either semantically related (thief/burglar) or had no obvious semantic relationship (farmer/burglar). Using this procedure, Ngo, Sargent, and Dopkins (2007) found that the false-alarm rate for the test word was higher with related than with unrelated primes. We expected to replicate this finding and that the mean of confidence ratings for test-word lures would be larger with related than with unrelated primes. We asked whether the standard deviation of confidence ratings for test-word lures would also be larger with related than with unrelated primes.

If we observed the latter result we were interested in how the UVSD and DPSD models would explain it. We assumed that the models would be best able to explain the result without additional parameters if performance was examined separately for related and unrelated primes. Assuming correspondence between distributions of confidence ratings and mnemonic evidence we reasoned as follows: To accommodate the result the UVSD model will posit that the standard deviation of the lure evidence distribution was larger with related than with unrelated primes. In contrast, the DPSD model will

accommodate the result by positing a lower rate of target recollection with related than with unrelated primes. Because the model attributes the size difference between the standard deviations of the target and lure distributions to the fact that recollection is more heavily involved in responses to targets than lures, the model will attribute the smaller size difference with related than with unrelated primes to the fact that less target recollection occurred with related than with unrelated primes. We found this interpretation problematic for the DPSD method of estimating the probability of target recollection, as we explain later.

Method

Participants

The participants were 195 students at the George Washington University.

Design

In each of the ten blocks of the experiment, a participant read a study list of 32 words, counted backwards for 3 s, and then responded to each of 32 pairs of words. To somewhat camouflage the relatedness manipulation, eight of the pairs for each block were filler pairs and only 24 of the pairs were experimental pairs. In each block the prime word and test word were semantically related in half (12) of the experimental pairs and had no obvious semantic relationship in half of the experimental pairs. The prime word and test word were never related in the filler pairs. The prime word was a target with a probability of .5 in the experimental pairs. The test word was a target in exactly half of the experimental pairs. The prime word and test word were targets with a probability of .5 in the filler pairs. In sum, across experimental and filler pairs, the prime word and test word were targets and lures, respectively, approximately 50% of the time, and the prime word and test word were related 37% of the time. Thus relationship (Related/Unrelated) and Test Word (Target/Lure) were manipulated within participants. Crossing these two independent variables produced an experimental design with four cells. In each of the ten blocks six experimental pairs were presented for each of the four cells.

Materials

The experimental pairs were generated from a set of 240 common noun triples. The first and second words of each triple were synonyms. The third word had no obvious semantic relationship to the first word. The rest of the words used in the experiment were sampled from a set of 620 common nouns. The materials for a given participant were created as

follows: First, the 240 triples were randomly assigned to the ten blocks, 24 triples to a block. Given the triples that had been assigned to a block, the materials for the block were created as follows: The triples for the block were randomly assigned to the four cells of the design, six triples to a cell. For triples assigned to the Related/Target and Related/Lure cells, the second and first words of the triple became the prime word and test word of an experimental pair. For triples assigned to the Unrelated/Target and Unrelated/Lure cells, the third and first words of the triple became the prime word and test word of an experimental pair. The test words of pairs assigned to the Related and Unrelated/Target cells were placed in the study list. The prime words of all pairs were placed in the study list with a probability of .5. The remaining words of the study list were then sampled without replacement from the set of 620 common nouns. Half of these words were placed as prime words or test words in filler pairs. The remaining words of the filler pairs were then sampled without replacement from the set of 620 common nouns.

Procedure

The study list for each block was presented with each successive word appearing alone in the middle of the screen for 1,600 ms and being followed by an inter-word interval of 400 ms. During the test phase the following occurred for each pair: A message appeared announcing the next pair. After 1,000 ms, this message disappeared and the prime word appeared. This word remained on the screen until the participant responded. The prime word then disappeared and the test word appeared. This word remained on the screen until the participant responded. The participant was then asked to indicate confidence in the response to the test word, using the three-point scale: 1 = Very confident; 2 = Somewhat confident; 3 = Not sure. The participant's initial instructions emphasized that the required confidence rating pertained only to the test word.

Results

The false-alarm rate (FAR) was greater in the Related than in the Unrelated condition [$t(194) = 10.10$] (see Table 1). The hit rate did not vary with relatedness. When we recoded the confidence ratings (CR) for positive and negative responses onto a single scale the mean [$t(194) = 10.50, p < .0001$] and standard deviation [$t(194) = 7.82$] were greater in the Related/Lure than in the Unrelated/Lure cell. The mean and standard deviation of the confidence ratings did not vary with relatedness in the Target condition.

The effects of relatedness on the false-alarm rate and the mean Lure confidence rating were greater in the first than the last five blocks of the experiment [FAR: $F(1,194) = 10.45,$

$MSe = .005$; CR: $F(1,194) = 16.31, MSe = .066$]. This probably reflects an adaptation effect. The effects of relatedness on the false-alarm rate and the mean Lure confidence rating were greater when the response to the prime word was 'no' as opposed to 'yes' [FAR: $F(1,194) = 10.65, MSe = .005$; CR: $F(1,194) = 18.67, MSe = .089$]. These findings may be related to the findings that the false-alarm rate and mean Lure confidence rating were higher when the response to the prime word was 'no' as opposed to 'yes' [FAR: $F(1,194) = 80.95, MSe = .015$; CR: $F(1,194) = 196.46, MSe = .201$]. For a previous report of this pattern – essentially a bias to make a response complementary to the previous response – see Dopkins, Sargent, and Ngo (2010). In the present context the bias may have been important in that the relatedness effect was more strongly expressed when the false-alarm rate and confidence rating were higher. To ensure that our results did not reflect a strategic effect of some sort we re-did the experiment asking participants to simply read the prime word without responding to it. All of our results replicated under these conditions: FAR, $t(114) = 7.22$; CR standard deviation, $t(114) = 6.31$; z-ROC functions, $t(114) = 2.08$; DPSD estimate of recollection lower in Related than Unrelated condition, $t(114) = 2.34$ (the latter two results are described later).

Collapsing across participants, we plotted ROC and z-ROC curves for the Related and Unrelated conditions (see Figs. 1 and 2). The z-ROC data were well fit with linear functions when hits were plotted as a function of false alarms (Related: $R^2 = .9983$, Unrelated: $R^2 = .9985$). Similar results were obtained when false alarms were plotted as a function of hits. In z-ROC functions for the individual participants, plotting hits as a function of false alarms, the average linear coefficient was larger for the Related (.78) than for the Unrelated (.71) condition [$t(194) = 3.63, p < .001$]. The effect of relatedness on the linear coefficient did not vary as a function of blocks or prime-word response. Consistent with the other effects of relatedness, the average linear coefficient was essentially 1 (1.02) when Related hits were plotted as a z-ROC function of Unrelated hits and less than 1 (.93) when Related false alarms were plotted as a z-ROC function of Unrelated false alarms [$t(194) = 7.82, p < .0001$].

We fit Gaussian instantiations of the UVSD and DPSD models to the data for each participant. The data for each fit comprised the numbers of assignments in the first five confidence categories for the four cells of the design. To find the best-fitting values of the parameters for each model we minimized the G statistic. Once we had found the best-fitting parameter values, we conducted a more focused test of the two models, asking whether they could account for the critical relatedness effect – the finding that the variability of confidence ratings was greater for related than for unrelated lures. We conducted this further test because the relatedness effect was small enough relative to other effects in the data that the G statistic would not be expected to reflect whether the effect

Table 1 Response rate and confidence rating data

	Target		Lure	
	Related	Unrelated	Related	Unrelated
Probability of positive response	.69 (.01)	.68 (.01)	.24 (.01)	.18 (.01)
Mean confidence rating	4.44 (.04)	4.42 (.04)	2.57 (.04)	2.33 (.04)
Standard deviation of confidence rating	1.63 (.02)	1.67 (.02)	1.44 (.03)	1.31 (.03)

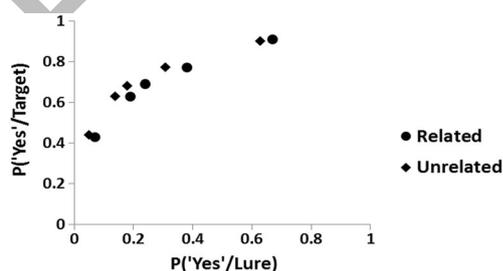
Note. Confidence ratings were recoded onto a single scale as follows: Very confident yes: 6, Somewhat confident yes: 5; Not sure yes: 4; Not sure no: 3; Somewhat confident no: 2; Very confident no: 1

Standard errors of mean are shown in parentheses

was captured by the model (Schunn & Wallach, 2005). To get a sense of the size of the relatedness effect relative to other effects in the data, consider the results of a repeated-measures ANOVA that used as data points the numbers of assignments for each participant in the first five confidence categories for the four cells of the design. Whereas the effect of confidence category produced an η_p^2 of .56, the effect of relatedness produced an η_p^2 of .08 (Bakeman, 2005).

To find out whether the two models could account for the relatedness effect, we asked whether they could account for the fact that the average linear z-ROC coefficient was larger in the Related than the Unrelated condition. Specifically, we used best-fitting parameter values for each participant and each model to generate numbers of assignments in the six confidence categories for the Related and Unrelated conditions and then fit z-ROC functions to the model-predicted data as in the analysis of the actual confidence data described earlier. The goal was to find out whether, as in the actual confidence data, the average linear z-ROC coefficient for the model-predicted data was larger for the Related than the Unrelated condition.

In an attempt to account for the data without added parameters we first fit each model separately to the data for the Related and Unrelated conditions. We fit models comprising the parameters shown in Table 2. The fit of both models was good (see Table 3). Crucially, both models could account for the relatedness effect: the average linear z-ROC coefficient for the model-predicted data was larger in the Related than in the Unrelated condition [UVSD: $t(194) = 7.50$, $p < .0001$, DPSD: $t(194) = 10.00$,

**Fig. 1** Aggregate receiver operating characteristic (ROC) functions based on rating data for Related and Unrelated conditions

$p < .0001$]. In the fits to the UVSD model, the value of σ_L was larger, on average, for the Related than for the Unrelated condition [$t(194) = 2.11$, $p < .05$]. In the fits to the DPSD model, the value of R_T was smaller, on average, for the Related than for the Unrelated condition [$t(194) = 2.84$, $p < .005$].

We next attempted to fit the data for the Related and Unrelated conditions together without added parameters (see Table 2). The fit for neither of the models was good (see Table 3). Crucially, neither model could account for the relatedness effect. To improve the fits we augmented each model with one additional parameter (see Table 2). Under the UVSD model we assumed that recognition judgments in the Related and Unrelated conditions were carried out in terms of distributions with different standard deviations. Under the DPSD model we assumed that recognition judgments in the Related condition involved false recollection of lures. The fit for the augmented models was better than the fit for the standard models (see Table 3). Crucially, both models could account for the relatedness effect. The average linear z-ROC coefficient for the model-predicted data was larger in the Related than in the Unrelated condition, UVSD: $t(194) = 3.88$; DPSD: $t(194) = 3.65$.

In the individual fits to the augmented UVSD model, the value of σ_{RL} was larger, on average, than the value of σ_{UL} [$t(194) = 13.16$, $p < .0001$]. In addition, the value of D_U was larger than the value of D_R [$t(194) = 3.35$, $p < .001$]. In the fits to the DPSD model, the value of D_U was larger than the value of D_R [$t(194) = 2.89$, $p < .005$]. The average level of false recollection was .02. Reassuringly, the same level of false

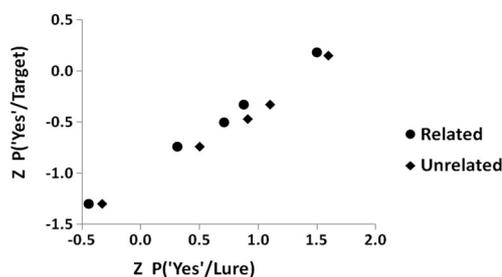
**Fig. 2** Aggregate z-ROC functions based on rating data for Related and Unrelated conditions

Table 2 Parameters of UVSD and DPSD models and Best-fitting Values[†]

		Related	Unrelated
UVSD - Related and Unrelated conditions fit separately			
D	Difference between target and lure means	1.04	1.18
σ_L	Standard deviation of lure distribution	.76	.71
Standard deviation of target distribution = 1			
Standard UVSD – Related and Unrelated conditions fit together			
D_R	Difference between target* and Related lure means	.91	
D_U	Difference between target* and Unrelated lure means	1.05	
σ_L	Standard deviation of lure distributions	.67	
Standard deviation of target* distribution = 1			
Augmented UVSD – Related and Unrelated conditions fit together			
D_R	Difference between target* and Related lure means	.94	
D_U	Difference between target* and Unrelated lure means	.99	
σ_{RL}	Standard deviation of Related lure distribution	.74	
σ_{UL}	Standard deviation of Unrelated lure distribution	.62	
Standard deviation of target* distribution = 1			
		Related	Unrelated
DPSD - Related and Unrelated conditions fit separately			
D	Difference between target and lure means	1.05	1.21
R_T	Probability of target recollection	.19	.24
Standard deviations of target and lure distributions = 1			
Standard DPSD – Related and Unrelated conditions fit together			
D_R	Difference between target* and Related lure means	.97	
D_U	Difference between target* and Unrelated lure means	1.15	
R_T	Probability of target recollection	.22	
Standard deviations of target and lure distributions = 1			
Augmented DPSD – Related and Unrelated conditions fit together			
D_R	Difference between target* and Related lure means	1.00	
D_U	Difference between target* and Unrelated lure means	1.11	
R_T	Probability of target recollection	.24	
R_L	Probability of false (lure) recollection	.02	
Standard deviations of target* and lure distributions = 1			

[†] Each model also included five parameters that articulated the confidence categories

* Related and Unrelated conditions were assumed to share the same target distribution. Mean of target distribution = 0

recollection was estimated when the DPSD model was fit to the data for Related and Unrelated lures.

Table 3 Fits for UVSD and DPSD models

UVSD		DPSD	
Mean G	Mean R^2	Mean G	Mean R^2
Related and Unrelated conditions fit separately			
2.76	.97	2.91	.97
Related and Unrelated conditions fit together – Basic model			
11.78	.88	13.56	.89
Related and Unrelated conditions fit together – Augmented model			
10.95	.92	11.17	.92

As a control we also fit the data with a model in which the criterion cut-points but not the memory parameters were allowed to vary between the two conditions. This model could not account for the critical relatedness effect, the average linear z-ROC coefficient for the model-predicted data was not larger in the Related than in the Unrelated condition, $t(194) = 1.60, p > .05$.

Discussion

In an item recognition task semantic priming increased the false-alarm rate and the mean of confidence ratings for lures. Such priming also increased the standard deviation of the lure

rating distribution and the z-ROC slope for the recognition judgments. The latter results suggest that the priming increased the standard deviation of the lure evidence distribution. This conclusion follows somewhat tenuously from the rating data because the ratings were not necessarily linearly related to the underlying evidence values. The conclusion follows more convincingly from the z-ROC data under the assumption that the evidence distributions for the recognition process were Gaussian (Rouder, Pratte, & Morey, 2010). Previous demonstrations of alignment between rating data and z-ROC data support this assumption (Mickes, Wixted, & Wais, 2007; Wixted & Mickes, 2010). We must, of course, be aware that z-ROC slope can vary with non-mnemonic factors. For example, z-ROC slope increased with increases in the degree of liberal bias in several previous studies (Ratcliff & Starns, 2009; Van Zandt, 2000). In the present case one can see how processing a related prime might have engendered a more liberal bias in processing the test word. However, the failure of the control model in which only the response criteria varied casts doubt on this interpretation. More generally, past results suggest that participants are reluctant to make trial-by-trial changes in decision criteria (Stretch & Wixted, 1998).

In terms of theory, our results pose problems for the UVSD and DPSD models. The UVSD conclusion that the standard deviation of the lure evidence distribution was larger for related than unrelated primes seems plausible but is completely ad hoc. In addition, some may question the plausibility of the DPSD conclusion that target recollection probability was lower following related than unrelated primes when the data for related and unrelated primes were fit separately, without added parameters. Further, this conclusion suggests problems generally with DPSD estimates of target recollection probability – suggesting, specifically, that such estimates will decrease with increases in the standard deviation of the lure confidence/evidence distribution. The reason for this was outlined in the Introduction.

The outcome when the data for related and unrelated primes were fit together suggests a means whereby the DPSD model might correct for differences in the lure confidence/evidence standard deviation; such differences might be attributed to differences in false recollection probability. To demonstrate how such a correction might work we again fit the data for related primes assuming false recollection at the level estimated in our original joint fit of the data for the related and unrelated primes. On average, the estimated target recollection probability (.26) in the resulting fits differed from the estimated probability in the original fits of the related-prime data (fits that did not assume false recollection) (.19) [$t(194) = 6.71$] and was similar to the estimated probability in the original fits of the unrelated-prime data (.24). So this correction might work. However, a problem with this approach is that the size of a given lure confidence/evidence standard deviation

relative to other possible lure confidence/evidence standard deviations is often unspecified by context. Hence the model often has no way of correcting its estimates of target recollection probability.

To illustrate, targets and lures must be related to some degree in any recognition experiment. Furthermore, ROC data suggest that the lure confidence/evidence standard deviation will increase with increases in the degree of target-lure relatedness (Dopkins, 2016; Rotello, Macmillan, & Van Tassel, 2000). The present results suggest that, unless false recollection probability is included in the DPSD model for a given experiment and allowed to be larger the greater the target-lure relatedness for the experiment (and thus larger the greater the lure confidence/evidence standard deviation for the experiment), target recollection probability will be incorrectly estimated as being smaller the greater the target-lure relatedness for the experiment (Dopkins (2016) has confirmed this). But what is the referent by which one establishes the degree of target-lure relatedness in a given experiment and thereby estimates false recollection probability?

In sum our results suggest that the lure confidence/evidence standard deviation varied with semantic priming in our experiment. We have suggested that such variation may be problematic for the DPSD model.

Author note We are grateful for Jeff Starns for suggesting that we fit the data for the Related condition with and without a parameter for false recollection. This research was supported by MH066189.

References

- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, *37*, 379–384.
- Dopkins, S. (2016). The lure rating distribution varies with target-lure relatedness: Implications for estimates of recollection. Submitted.
- Dopkins, S., Sargent, J., & Ngo, C. T. (2010). The bias for a recognition judgment depends on the response emitted in a prior recognition judgment. *Memory*, *18*, 272–283.
- Jang, Y., Mickes, L., & Wixted, J. T. (2012). Three tests and three corrections: Comment on Koen and Yonelinas (2010). *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *38*, 513–523.
- Koen, J. D., & Yonelinas, A. P. (2010). Memory variability is due to the contribution of recollection and familiarity, not to encoding variability. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *36*, 1536–1542.
- Lipton, P., & Eichenbaum, H. (2012). Towards a functional organization of episodic memory in the medial temporal lobe. *Neuroscience and Biobehavioral Reviews*, *36*, 1597–1608.
- Mickes, L., Wixted, J. T., & Wais, P. E. (2007). A direct test of the unequal-variance signal detection model of recognition memory. *Psychonomic Bulletin & Review*, *14*, 858–865.
- Ngo, C. T., Sargent, J., & Dopkins, S. (2007). Level of discrimination for recognition judgments reduced following the recognition of semantically related words. *Journal of Memory and Language*, *57*, 415–436.

- Ratcliff, R., Sheu, C., & Gronlund, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review*, *99*, 518–535.
- Ratcliff, R., & Starns, J. J. (2009). Modeling confidence and response time in recognition memory. *Psychological Review*, *116*, 59–83.
- Rotello, C. M., Macmillan, N. A., & Van Tassel, G. (2000). Recall-to-reject in recognition: Evidence from ROC curves. *Journal of Memory and Language*, *43*, 67–88.
- Rouder, J. N., Pratte, M. S., & Morey, R. D. (2010). Latent mnemonic strengths are latent: A comment on Mickes, Wixted, and Wais (2007). *Psychonomic Bulletin & Review*, *17*, 427–435.
- Schoemaker, D., Gauthier, S., & Pruessner, J. C. (2014). Recollection and familiarity in aging individuals with mild cognitive impairment and Alzheimer's disease: A literature review. *Neuropsychology Review*, *24*, 313–331.
- Schunn, C. D., & Wallach, D. (2005). Evaluating goodness-of-fit in comparison of models to data. In W. Tack (Ed.), *Psychologie der Kognition: Reden and Vorträge anlässlich der Emeritierung von Werner Tack* (pp. 115–154). Saarbrücken: University of Saarland Press.
- Starns, J. J., Rotello, C. M., & Ratcliff, R. (2012). Mixing strong and weak targets provides no evidence against the unequal variance explanation of zROC slope: A comment on Koen and Yonelinas (2010). *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *38*, 793–801.
- Stretch, V., & Wixted, J. T. (1998). Decision rules for recognition memory confidence judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *24*, 1397–1410.
- Van Zandt, T. (2000). ROC curves and confidence judgments in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 582–600.
- Wixted, J. T. (2007). Dual-process theory and signal detection theory of recognition memory. *Psychological Review*, *114*, 152–176.
- Wixted, J. T., & Mickes, L. (2010). Useful scientific theories are useful: A reply to Rouder, Pratte, and Morey (2010). *Psychonomic Bulletin & Review*, *17*, 436–442.
- Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory: A review. *Psychological Bulletin*, *133*, 800–832.

FOR APPROVAL