BRIEF REPORTS

“Remembering” emotional words is based on response bias, not recollection

SONYA DOUGAL

New York University, New York, New York

AND

CAREN M. ROTELLO

University of Massachusetts, Amherst, Massachusetts

Recent studies have demonstrated that emotional stimuli result in a higher proportion of recognized items that are “remembered” (e.g., Kensinger & Corkin, 2003; Ochsner, 2000), leading to greater estimates of recollection by the dual-process model (Yonelinas, 1994). This result suggests that recognition judgments to emotional stimuli depend on a recollection process. We challenge this conclusion with receiver operating characteristic (ROC) curve data from two experiments. In both experiments, subjects studied neutral and emotional words. During the recognition test, subjects made old–new confidence ratings as well as remember–know judgments. Four models of remember–know judgments were fit to individual subjects’ data: two versions of a one-dimensional signal-detection-based model (Donaldson, 1996; Wixted & Stretch, 2004), the dual-process model (Yonelinas, 1994), and the two-dimensional signal-detection-based model known as STREAK (Rotello, Macmillan, & Reeder, 2004). Consistent with the literature, we found that emotion increases subjective reports of “remembering.” However, our ROC analyses and modeling work reveal that the effect is due to response bias differences rather than sensitivity change or use of a high-threshold recollection process.

Emotionally significant events are typically believed to be remembered with extraordinary accuracy. Indeed, the subjective experience of memory for emotional events is often more vivid and detail laden than memory for ordinary events. The present research asks whether the enhanced experience of recollection associated with recognition of emotional stimuli is the result of an increase in memory accuracy brought about by the contribution of recollection to recognition judgments.

One method for testing the belief that emotion enhances the recollection of details is to collect subjective reports of “remembering” during recognition (Tulving, 1985). In the remember–know paradigm, subjects encode a list of words and are given a recognition test that requires them to indicate whether they “remember” episodic details from having studied each word, or whether they “know” the word was studied despite the failure to recall any details. Recent studies indicate that the proportion of “remember” judgments to emotionally arousing stimuli is elevated in comparison with neutral stimuli, whereas “know” responses do not differ with emotion (Kensinger & Corkin, 2003; Ochsner, 2000; Sharot, Delgado, & Phelps, 2004). Because “remember” responses are thought to measure recollection of neutral stimuli (see, e.g., Yonelinas, 2002),

this finding has led to the claim that recognition memory for emotional stimuli depends largely on a recollective process (e.g., Kensinger & Corkin, 2003; Ochsner, 2000), such as that described by Yonelinas (1994) in his dual-process model of recognition.

A standard way of measuring accuracy is to calculate d' , the distance between the means of the memory strength distributions of studied items and lures. Several studies have asked whether recognition accuracy is enhanced for emotional stimuli, with mixed results. Although recognition accuracy is sometimes enhanced for emotional in comparison with neutral stimuli (see, e.g., Kensinger & Corkin, 2003, Experiments 1 and 3), other studies show no effect of emotion or a decrease in d' (e.g., Johansson, Mecklinger, & Treese, 2004; Kensinger & Corkin, 2003, Experiment 5; Maratos, Allen, & Rugg, 2000; Sharot et al., 2004; Windmann & Kutas, 2001). In contrast, our survey of the literature revealed that *all* studies show consistent changes in response bias with emotion. Unfortunately, when response bias differs across conditions, d' cannot be meaningfully compared unless the *old* and *new* distributions have the same variance. In most recognition experiments, the distribution of *old* item memory strengths is more variable than the *new* item distribution (e.g., Ratcliff,

S. Dougal, sonya.dougal@nyu.edu

Table 1
Mean Response Proportions and Summary Statistics for Each Stimulus Type
in Experiments 1A and 1B

	Dependent Measure	Experiment 1A			Experiment 1B		
		Neutral	Negative	Positive	Neutral	Negative	Positive
Old items	P ("old")	.57	.69	.53	.62	.76	.60
	P ("remember")	.33	.44	.27	.34	.42	.32
New items	P ("old")	.13	.34	.22	.19	.34	.21
	P ("remember")	.02	.08	.06	.06	.11	.06
Summary statistics	d'	1.54	1.02	0.95	1.31	1.21	1.17
	c	.55	-.02	.40	.31	-.19	.28

Sheu, & Gronlund, 1992). In that case, d' tends to overestimate true sensitivity for conservative criteria and to underestimate it for liberal criteria. Thus, an apparent sensitivity difference, measured by d' , may actually reflect only a change in response bias between conditions.

The solution to the dependence of d' and bias is to use receiver operating characteristic (ROC) curves to evaluate whether memory sensitivity, response bias, or both are affected by the emotionality of the stimuli. An ROC curve plots the hit rate against the false alarm rate as a function of bias. The curves allow clear evaluation of differences in bias—points that fall on a common theoretical curve—and differences in sensitivity—points that fall on distinct curves (see Macmillan & Creelman, 2005). Collecting ROC data has the added advantage that it constrains the fits of memory models that make distinct predictions about the processes underlying remember-know judgments (Rotello, Macmillan, Hicks, & Hautus, 2006; Rotello, Macmillan, & Reeder, 2004). A simple method for obtaining ROC data, used here, is to ask subjects to rate their confidence that each test item was studied.

We report two experiments in which subjects studied a series of neutral and emotional words. The subjects were then given a recognition test. They made their responses on a 6-point confidence scale, and supplemented each "old" response with a subjective report of the basis of that decision ("remember" or "know"). The goals of the study were to assess (1) whether emotional words resulted in increased memory sensitivity, (2) whether response bias changed as a function of emotion, and (3) whether emotion increased the use of a recollective process in recognition. Goals 1 and 2 were addressed with ROC data. Goal 3 was evaluated by fitting quantitative models of memory to the data that make distinct assumptions about the nature of "remembering."

EXPERIMENTS 1A AND 1B

The experiments differed in materials only.

Method

Subjects. Sixty New York University undergraduates participated in exchange for extra credit: 30 in Experiment 1A¹ and 30 in Experiment 1B.

Design. We used a 2×3 factorial design with two within-subjects factors: item status (*studied*, *new*) and emotionality of the words (negative and arousing, neutral, or positive and arousing).

Materials. Stimuli were selected from the ANEW pool of words (Bradley & Lang, 1999), for which self-reported arousal levels are correlated with physiological measures of arousal (Lang, Bradley, & Cuthbert, 1998). In Experiment 1A, there were 96 negative-arousing words (mean valence = 2.24; mean arousal = 6.63), 96 positive-arousing words (mean valence = 7.79; mean arousal = 6.48), and 96 neutral, nonarousing words (mean valence = 5.16; mean arousal = 4.15). Francis and Kučera (1982) word frequencies were equated across emotion. In Experiment 1A, semantic interrelatedness among the negative and positive words was matched using latent semantic analysis (Landauer, Foltz, & Laham, 1998). In Experiment 1B, we added an additional 96 neutral words to increase the semantic relatedness of the neutral words, thereby matching semantic relatedness among *all* classes of words, while retaining the other characteristics of the materials.

Procedure. The subjects were instructed to study 144 words (48 neutral, 48 negative, 48 positive) in Experiment 1A, or 192 words (96 neutral, 48 negative, 48 positive) in Experiment 1B. The words were shown one at a time for 3 sec each. Immediately following presentation of the study list, the subjects made recognition judgments for 288 (Experiment 1A) or 384 (Experiment 1B) test words (half *studied*, half *new*) on a scale of 1 (*sure old*) to 6 (*sure new*). For items judged to be "old," subjects used Rajaram's (1993) descriptions of remembering and knowing to make remember-know judgments. Study and test sequences were randomly ordered for each subject.

Results

We consider the data in three ways: (1) overall response proportions collapsed across confidence ratings, for comparison with the literature; (2) ROC curves, for clear evaluation of sensitivity and bias; (3) theoretical analysis of the data with four models that make different assumptions about the interpretation of "remember" and "know" judgments.

Behavioral data. Overall hit and false alarm rates, and the proportion of "remember" and "know" responses given to *old* and *new* words, are shown in the first four rows of Table 1. In Experiment 1A, the proportion of "old" judgments was higher for negative words ($M = .51$) than for neutral words ($M = .35$) and positive words ($M = .38$), regardless of whether the words had been studied [$F(2,58) = 38.45$, $MS_e = .01$, $p < .01$]. This liberal response bias for negative stimuli is consistently observed in the literature. Negative words were also more likely to be accompanied by the subjective experience of "remembering" ($M = .26$) than were neutral words ($M = .17$) or positive words ($M = .16$), regardless of whether the words had been studied [$F(2,58) = 15.24$, $MS_e = .01$, $p < .01$]. Both effects were replicated in Experiment 1B: The proportion

of “old” judgments was higher for negative words ($M = .55$) than for neutral words ($M = .40$) and positive words ($M = .41$), regardless of whether the words had been studied [$F(2,58) = 57.64$, $MS_e = .01$, $p < .01$]. In addition, negative words were more likely to be accompanied by the subjective experience of “remembering” ($M = .26$) than were neutral words ($M = .20$) or positive words ($M = .20$) [$F(2,58) = 13.90$, $MS_e = .01$, $p < .01$]. These findings replicate previous reports that emotion increases the proportion of “remember” judgments (e.g., Ochsner, 2000).

In Experiment 1A, recognition accuracy was also influenced by emotion. The influence of emotion on the proportion of “old” judgments was larger for *new* words than for *old* words [$F(2,58) = 13.11$, $MS_e = .01$, $p < .01$], resulting in lower memory sensitivity for the negative words (mean $d' = 1.02$) and positive words ($d' = 0.95$) than for the neutral words ($d' = 1.54$) [$F(2,58) = 21.70$, $MS_e = .14$, $p < .01$; see Table 1, row 5]. These data replicate Maratos et al. (2000), who found that emotion reduced recognition accuracy when the emotional words were more semantically interrelated than the neutral words. The response bias measure c was also lower for negative words ($M = -.02$) than for neutral words ($M = .55$) and positive words ($M = .40$) [$F(2,58) = 44.51$, $MS_e = .06$, $p < .01$; see Table 1, row 6]. In Experiment 1B, semantic relatedness was matched across all words, and we observed no effect of emotion on recognition accuracy ($F < 1.50$; see Windmann & Kutas, 2001). Thus, equating the semantic relatedness of the word types eliminated any effect of emotion on memory accuracy. We did replicate Experiment 1A's result that negative emotion resulted in a more liberal response bias. In Experiment 1B, c was lower for negative words ($M = -.19$) than for neutral words ($M = .31$) and positive words ($M = .28$) [$F(2,58) = 63.03$, $MS_e = .04$, $p < .01$]. Because response bias differed across conditions, we turn to the ROC data for a more appropriate evaluation of memory sensitivity.

ROC data. ROC curves were created by plotting hit rates against false alarm rates as a function of confidence. The highest confidence “old” decisions constitute the leftmost point in ROC space; decreasing confidence is reflected in the points to the right. Three aspects of the ROC are relevant: (1) all points on the same ROC curve reflect the same memory sensitivity; (2) higher sensitivity is indicated by curves that fall toward the upper-left portion of the space, where the hit rate is higher relative to the false alarm rate; and (3) different response bias is indicated by a change in the location of points along the same curve. If two points fall on the same curve, the point that is closer to (1, 1) reflects a more liberal response bias, because it reflects a higher hit rate as well as a higher false alarm rate than a point closer to (0, 0).

Figure 1 shows the old–new ROC points for each stimulus type, on the basis of group data from Experiment 1A (top row) and 1B (bottom row). Two main effects of emotion are apparent in these data. First, memory sensitivity is higher for the neutral than for the emotional words in Experiment 1A, as can be seen by the higher position of the ROC curve for neutral items. This sensitivity ef-

fect disappears in Experiment 1B, in which the neutral and emotional stimuli were matched on semantic relatedness. Second, response bias was more liberal for the negative than for the neutral and positive words in both experiments, as can be seen by the higher position of the negative point in comparison with the neutral and positive points at each level of confidence. The asymmetrical shape of these old–new ROC curves is consistent with both the dual-process model (Yonelinas, 1994) and an unequal-variance signal-detection model (Macmillan & Creelman, 2005). Next, we fit different models of remember–know judgments to individual subjects' ROC curves and their remember–know response rates at each confidence level to determine whether the increased “old” and “remember” response rates to negative words are a function of more detailed recollection for emotional than for neutral stimuli.

Model-based analyses. There are four dominant quantitative models of remember–know judgments that make distinct assumptions about the meaning of “remember” responses (see Rotello et al., 2006, for model details and equations). We fit these models to our data with the goal of appropriately interpreting the increased “remember” rates to the negative stimuli: Do the “remember” judgments reflect increased recollection of those stimuli, or not?

In both of the one-dimensional signal-detection models, “remember” responses simply reflect item familiarity; they do not result from recollection. In these models, subjects set a criterion level of strength, above which, items are called “old.” Remember–know judgments are made by setting a second criterion, above which, items are said to be “remembered,” and below which, they are “known” to be old (Donaldson, 1996). The two versions of this model differ only in their assumptions about the stability of the decision criteria over trials: The fixed version assumes that there is no variability in the locations of the criteria (Donaldson, 1996), whereas the variable-criterion version assumes that the remember–know criterion is variable (Wixted & Stretch, 2004). Figure 2A summarizes these models graphically.

The second model we fit to our data is the dual-process model first proposed by Yonelinas (1994) and later extended to the remember–know paradigm (Yonelinas, 2001; Yonelinas, Dobbins, Szymanski, Dhaliwal, & King, 1996). In the dual-process model, “remember” and “know” judgments reflect distinct underlying memory processes. “Remember” judgments reflect the contribution of a high-threshold recollective process in which recollection is assumed to result in high-confidence “old” responses to *studied* items only. Because *new* items cannot be recollected, “remember” responses should predominately occur after “sure old” judgments for *studied* words. In contrast, “know” judgments can occur with any level of confidence, and reflect the contribution of an equal-variance signal-detection process, as depicted in Figure 2B.

Finally, we fit a two-dimensional signal-detection-based model known as STREAK (Rotello et al., 2004) to the data. STREAK (depicted in Figure 2C) assumes that old–new recognition judgments are based on the sum of two kinds of memory strength, global and specific,

whereas remember-know judgments are based on the difference of those strengths. Thus, STREAK shares the view of the one-dimensional model that remember-know judgments reflect different amounts of the same underlying memory information.

Each condition in this experiment provided a matrix of 18 response frequencies (6 confidence ratings $\times$ 2 types of items [*old* or *new*], where 3 of the ratings were followed by 2 types of judgment [remember or know]) for each subject. Each model was fit to the full pattern of “remember,” “know,” and “new” responses across confidence levels (i.e., for each model, we generated the predicted probability of each type of response at each confidence level, then minimized the differences between observed and predicted response probabilities). The best-fitting parameter values were found by using maximum-likelihood estimation (see Rotello et al., 2006). The models differ in their number of free parameters

and are not nested, so we used Akaike’s information criterion (AIC; Akaike, 1973) to compare the model fits while adjusting for the number of parameters. We fit all conditions simultaneously (though all criteria, sensitivity, and other parameters were free to vary across conditions), which resulted in a single AIC value for each subject-model combination. Table 2 shows the mean AIC value for each model.

Forty-one subjects’ data (19 in Experiment 1A; 22 in Experiment 1B) were best described by the one-dimensional model (either fixed [14 subjects in Experiment 1A; 10 in Experiment 1B] or variable criterion [5 subjects in Experiment 1A; 12 in Experiment 1B]); 19 subjects’ data (11 in Experiment 1A; 8 in Experiment 1B) were best fit by the dual-process model. STREAK did not provide the best fit for any subject, probably because the “remember” response rates were correlated with confidence level in our data, contrary to the model’s predictions; thus, we will not

Figure 1. Receiver operating characteristic (ROC) data. The upper panels show data from Experiment 1A (both panels show the same data); the lower panels report results of Experiment 1B (both panels show the same data). The filled circles denote the observed data from the neutral condition, the minus signs denote the observed points in the negative condition, and the plus signs indicate the data from the positive condition. Best-fitting model predictions are also superimposed: Dual-process fits are shown on the left and variable-criterion one-dimensional model fits on the right. Fixed-criterion one-dimensional model fits are similar to the variable-criterion predictions. Solid curves are model predictions for the neutral condition, dotted curves are for the negative condition, and dashed curves are for the positive condition.

Figure 2. How old–new response bias operates according to three models of remember–know judgments. (A) The one-dimensional signal-detection model. *Old* and *new* items vary in their average memory strength, and two kinds of criteria are used to determine responses: The remember–know criterion divides “remember” and “knows,” and the old–new criteria determine rating responses. (B) Dual-process model. If recollection occurs, a “remember” response is made; otherwise a judgment is based on familiarity. (C) STREAK. *Old* and *new* items differ in both specific (d_y) and global (d_x) strength. Circles represent equal-likelihood contours from bivariate distributions; the upper circle is the *old* distribution and the lower is the *new* distribution. The old–new decision bound distinguishes “old” from “new” responses on the basis of a weighted sum of the two axes, and rating responses are determined by bounds parallel to it. An orthogonal bound distinguishes “remember” from “know” responses on the basis of a weighted difference. From “Interpreting the Effects of Response Bias on Remember–Know Judgments Using Signal Detection and Threshold Models,” by C. M. Rotello, N. A. Macmillan, J. L. Hicks, and M. J. Hautus, *Memory & Cognition*, 34, p. 1599. Copyright 2006 by Psychonomic Society, Inc. Adapted with permission.

consider it further. In sum, for the majority of our subjects (68%), there was no evidence that “remember” responses reflected the retrieval of episodic details.

Although our conclusions are based on individual subjects’ fits, we also fit the group data so that we could provide a graphical indication of the model fits. In Figure 1, the best-fitting functions of the dual-process model (panels A and C) and the variable-criterion one-dimensional model (panels B and D) are superimposed on the group ROCs.² The top row shows Experiment 1A; the bottom shows Experiment 1B. Notice that the dual-process model systematically misses the conservative points on the ROCs. In addition, the cumulative group “remember” response rates are shown for each condition as a function of confidence level in Figure 3A (Experiment 1A) and 3B (Experiment 1B), along with the functions predicted by the dual-process and one-dimensional models. Notice that the observed “remember” response rates rise as the response criterion becomes more liberal: This pattern indicates that subjects spread their “remember” judgments over confidence levels. The one-dimensional model predicts that the “remember” response rates will rise as criterion becomes more liberal (the model predictions have a positive slope with confidence), but the dual-process

model does not (the model predictions are flat across confidence), because it assumes that essentially all “remember” responses follow “sure old” judgments.

Finally, the individual subjects’ parameter values were subjected to ANOVA. We found that response bias shifts were the primary influence of emotion for all models. The effects of emotion on the parameter values (shown in Table 3) are consistent with the behavioral effects seen in Table 1 in reflecting clear bias changes; sensitivity differences occurred only in Experiment 1A, in which the emotional words were more interrelated than the neutral words.

Discussion

Our results indicate that negative-arousing words consistently elicit more “old” judgments on a recognition test than do neutral or positive-arousing words. These “old” judgments to negative stimuli are associated with a greater probability of the subjective experience of “remembering” than are neutral or positive words. However, there was no effect of emotion on memory sensitivity when the semantic similarity of the neutral and emotional stimuli was equated (Experiment 1B). Instead, negative emotional arousal consistently resulted in a more lenient response bias. Finally, our modeling analysis indicated

Table 2
AIC Values, Averaged Over Subjects, for Each Model and Both Experiments

Experiment	Model			
	One-Dimensional, Fixed Criteria	One-Dimensional, Variable Criteria	Dual Process	STREAK
1A	858.51	863.63	908.47	992.50
1B	1,254.73	1,243.64	1,345.68	1,419.89

Note—Lower values of AIC indicate better fit. Values are larger in Experiment 1B because there were more data points in that experiment.

Figure 3. Group “remember” response rates as a function of (cumulative) confidence level in Experiment 1A (panel A) and Experiment 1B (panel B). Filled squares denote the observed data; superimposed functions indicate the predictions of the dual-process model (solid line), one-dimensional model with fixed criteria (short dashes), and one-dimensional model with variable remember-know criterion (long dashes), assuming the best-fitting parameter values in each case.

that “remember” judgments to emotional stimuli do *not* reflect use of a recollective process or retrieval of episodic details.

Implications for models of remember-know judgments. The one-dimensional signal-detection model provides the best general account of our data. The individual model fits indicated that this model most accurately de-

scribed the data for 68% of our subjects. On this view, subjects respond more liberally to negative stimuli than to neutral and positive ones, but do not show enhanced accuracy to those items. Moreover, according to the one-dimensional model, the subjective experience of “remembering” the emotional items was also simply an effect of response bias. These results with emotional stimuli are consistent with prior conclusions on recognition judgments more generally (Rotello et al., 2006).

The dual-process model provides a relatively poor description of these data, accounting for only 32% of the subjects’ data. This is likely because the dual-process model predicts that “remember” judgments will occur predominately at the highest level of confidence, whereas our subjects used a range of confidence levels to report “remembering” (see also Rotello, Macmillan, Reeder, & Wong, 2005). We conclude that emotion uniformly changes response bias and that “remember” judgments to emotional stimuli are made on the basis of item familiarity, not recollection. These behavioral findings are fundamentally inconsistent with the notion that a high-threshold recollective process contributes to the recognition of emotional stimuli.

Implications for the emotion literature. The belief that emotion enhances memory accuracy is predominant in the literature and receives some empirical support from previous studies of free recall (e.g., Doerksen & Shimamura, 2001; Phelps, LaBar, & Spencer, 1997).³ However, examination of studies of recognition of emotional stimuli indicates that support for this belief is weak. Some studies do find that emotion enhances recognition accuracy (Kensinger & Corkin, 2003; Ochsner, 2000). Yet others find no difference in memory accuracy between neutral and emotional stimuli (Comblain, D’Argembeau, Van der Linden, & Aldenhoff, 2004; Leiphart, Rosenfeld, & Gabrieli, 1993; Pesta, Murphy, & Sanders, 2001; Windmann & Kutas, 2001), or observe that recognition accuracy is *lower* for negative words than for neutral words (Maratos et al., 2000).

Our data suggest that inadequate control over the semantic similarity of the emotional and neutral words may contribute to these mixed results. In our data, the emotionally arousing words were associated with lower levels of recognition accuracy when they were more semantically

Table 3
Mean Parameter Values for the Best-Fitting Model in Experiment 1A (One-Dimensional Fixed) and Experiment 1B (One-Dimensional Variable Criterion)

Parameter	Experiment 1A			Experiment 1B		
	Neutral	Negative	Positive	Neutral	Negative	Positive
$zROC$ slope (a, b)	.59	.70	.78	.66	.88	.74
d'_1 (a)	1.98	1.48	1.12	1.74	1.51	1.45
c_1 (most liberal)	-1.33	-2.29	-1.96	-1.11	-0.81	-2.03
c_2 (a, b)	0.42	-0.37	-0.46	0.21	-0.07	0.07
c_3 (a, b)	1.50	0.58	0.95	0.97	0.45	0.89
c_4 (a, b)	2.15	1.09	1.44	1.56	1.08	1.52
c_5 (most conservative) (a, b)	3.16	1.62	1.86	2.14	1.65	2.32
c_r	2.44	1.80	2.15	1.72	1.31	1.97
t	n/a	n/a	n/a	4.64	7.54	6.97

Note— d'_1 is the distance between the *old* and *new* distributions measured in units of the *new* distribution; c_1 – c_5 are old–new decision criteria; c_r is the mean of the remember-know criterion; t is the standard deviation of the remember-know criterion location. Statistically reliable effects of emotion are indicated by a for Experiment 1A and b for Experiment 1B.

interrelated than the neutral words were. However, when semantic similarity was matched between the neutral and emotional words, there was a difference in response bias, but not in recognition accuracy. Indeed, Windmann and Kutas (2001) also matched the semantic relatedness of their emotional and neutral words and reported that emotion only changed response bias. A similar suggestion has been made about the effects of emotion on recall (Phelps et al., 1998).

In conclusion, emotion does not enhance recognition accuracy when semantic relatedness is controlled across the neutral and emotional stimuli. However, emotion does have two consistent effects on recognition. First, negative arousal results in more liberal response bias. Second, emotion enhances the subjective experience of recollection as measured by the remember-know procedure. Our modeling results suggest that the enhanced experience of recollection for emotional stimuli is based only on item familiarity, not true recollection. Taken together, these findings indicate that emotion merely creates the illusion of recollection in recognition memory.

AUTHOR NOTE

S.D. was supported by Ruth Kirschstein postdoctoral fellowship HRD 0332898 from the National Institutes of Health, and C.M.R. was supported by Grant R01 MH60274 from the National Institutes of Health. We thank John Wixted, Andy Yonelinas, and an anonymous reviewer for their insightful comments on the manuscript. Correspondence concerning this article should be addressed to S. Dougal, New York University, 6 Washington Place, Room 873, New York, NY 10003 (e-mail: sonya.dougal@nyu.edu).

REFERENCES

- AKAIKE, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrox & F. Caski (Eds.), *Second international symposium on information theory* (pp. 267-281). Budapest: Akadémiai Kiadó.
- BRADLEY, M. M., & LANG, P. J. (1999). *Affective norms for English words (ANEW): Stimuli, instruction manual and affective ratings* (Tech. Rep. C-1). Gainesville: University of Florida, Center for Research in Psychophysiology.
- COMBLAIN, C., D'ARMBEAU, A., VAN DER LINDEN, M., & ALDENHOFF, L. (2004). The effect of ageing on the recollection of emotional and neutral pictures. *Memory*, **12**, 673-684.
- DOERKSEN, S., & SHIMAMURA, A. P. (2001). Source memory enhancement for emotional words. *Emotion*, **1**, 5-11.
- DONALDSON, W. (1996). The role of decision processes in remembering and knowing. *Memory & Cognition*, **24**, 523-533.
- FRANCIS, W. N., & KUČERA, H. (1982). *Frequency analysis of English usage*. Boston: Houghton Mifflin.
- JOHANSSON, M., MECKLINGER, A., & TREESE, A. C. (2004). Recognition memory for emotional and neutral faces: An event-related potential study. *Journal of Cognitive Neuroscience*, **16**, 1840-1853.
- KENSINGER, E. A., & CORKIN, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition*, **31**, 1169-1180.
- LANDAUER, T. K., FOLTZ, P. W., & LAHAM, D. (1998). Introduction to latent semantic analysis. *Discourse Processes*, **25**, 259-284.
- LANG, P. J., BRADLEY, M. M., & CUTHBERT, B. N. (1998). Emotion and motivation: Measuring affective perception. *Journal of Clinical Neuropsychology*, **15**, 397-408.
- LEIPHART, J., ROSENFELD, P., & GABRIELI, J. D. (1993). Event-related potential correlates of implicit priming and explicit memory tasks. *International Journal of Psychophysiology*, **15**, 197-206.
- MACMILLAN, N. A., & CREELMAN, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.
- MARATOS, E. J., ALLEN, K., & RUGG, M. D. (2000). Recognition memory for emotionally negative and neutral words: An ERP study. *Neuropsychologia*, **38**, 1452-1465.
- OCHSNER, K. N. (2000). Are affective events richly recollected or simply familiar? The experience and process of recognizing feelings past. *Journal of Experimental Psychology: General*, **129**, 242-261.
- PESTA, B. J., MURPHY, M. D., & SANDERS, R. E. (2001). Are emotionally charged lures immune to false memory? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 328-338.
- PHELPS, E. A., LABAR, K. S., ANDERSON, A. K., O'CONNOR, K. J., FULBRIGHT, R. K., & SPENCER, D. D. (1998). Specifying the contributions of the human amygdala to emotional memory: A case study. *Neurocase*, **4**, 527-540.
- PHELPS, E. A., LABAR, K. S., & SPENCER, D. D. (1997). Memory for emotional words following unilateral temporal lobectomy. *Brain & Cognition*, **35**, 85-109.
- RAJARAM, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition*, **21**, 89-102.
- RATCLIFF, R., SHEU, C.-F., & GRONLUND, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review*, **99**, 518-535.
- ROTELLO, C. M., MACMILLAN, N. A., HICKS, J. L., & HAUTUS, M. J. (2006). Interpreting the effects of response bias on remember-know judgments using signal detection and threshold models. *Memory & Cognition*, **34**, 1598-1614.
- ROTELLO, C. M., MACMILLAN, N. A., & REEDER, J. A. (2004). Sum-difference theory of remembering and knowing: A two-dimensional signal detection model. *Psychological Review*, **111**, 588-616.
- ROTELLO, C. M., MACMILLAN, N. A., REEDER, J. A., & WONG, M. (2005). The remember response: Subject to bias, graded, and not a process-pure indicator of recollection. *Psychonomic Bulletin & Review*, **12**, 865-873.
- SHAROT, T., DELGADO, M. R., & PHELPS, E. A. (2004). How emotion enhances the feeling of remembering. *Nature Neuroscience*, **12**, 1376-1380.
- TULVING, E. (1985). Memory and consciousness. *Canadian Psychology*, **26**, 1-12.
- WINDMANN, S., & KUTAS, M. (2001). Electrophysiological correlates of emotion-induced recognition bias. *Journal of Cognitive Neuroscience*, **13**, 577-592.
- 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, & Cognition*, **20**, 1341-1354.
- 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 & Language*, **46**, 441-517.
- 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 & Cognition*, **5**, 418-441.

NOTES

1. Three additional subjects' data were discarded because they responded only with extreme confidence ratings, which did not allow for stable parameter estimates of the models.
2. The group data in Experiment 1A were better fit by the variable-criterion model than the fixed version. Although most individual subjects showed stable remember-know criteria, averaging the data makes the remember-know criterion appear unstable. This result is analogous to the observation that averaging a set of step functions can result in a function that appears to show continuous change.
3. An important limitation in interpreting the recall studies is that intrusion data are typically not reported. Thus, increased recall of emotional stimuli may reflect a reporting bias rather than increased memory accuracy.

(Manuscript received April 12, 2006;
revision accepted for publication June 7, 2006.)